

Development of a real-time all-in-focus microscope

— WYSIWYG in the micro-world —

Kohtaro Ohba

[Translation from *Synthesiology*, Vol.2, No.4, p.264-275 (2009)]

In this paper, our struggle to realize a high-speed digital processed microscopic observational system for tele-micro-operation with a dynamic focusing system and a high-speed digital-processing system using the “depth from focus” criteria is reported. To realize the system, each functional element and its system configuration had been deeply discussed not only in the academic society but also with several companies, there were many trials and errors, and the final product system had been developed after several trials.

Keywords : Real-time, microscope, all-in-focus

1 Technological issue to be realized

Recently, there have been increase in industrial enthusiasm in micro- and nano-technologies, and there is also an increase in the demands for systems that can observe the microenvironment and systems that allow manipulators to operate the microenvironment. The examples of observation systems are optical and electron microscopes. The optical microscopes are widely used to observe micro-size objects that do not surpass the optical limit; i.e. sub micro meter, and lack of the depth information. In biological usage, the products demanded in the micro-size market include ones that allow cell and DNA manipulation while looking at the optical microscope images. In industrial usage, there are demands for observing and inspecting both the wire-bonded chip surface and bonding surface at the same time in LSI product inspection.

In such microscopic operation, it is necessary to conduct manipulations while sensing the three dimensional position of the actual micro-object. The major differences between an ordinary environment and a microenvironment are, in the latter, (1) the viscosity due to the van der Waals force rather than the weight of the object due to physical scaling effect cannot be neglected, and at the same time (2) as we approach the optical limit of the optical microscope, the range in the depth of the visible object (depth of field) becomes extremely small due to optical scaling effect. In this research, the issue of (2) optical scaling effect is addressed.

As a result of the optical scaling effect, in an optical system with shallow depth of field as in a microscope image, the objects with different depths are unfocused when an object with certain depth is set in focus, as shown in Fig. 1. Therefore, for product inspection in many cases, several images are shot at varying focal distances while moving the focal distance, and then these images are processed. Therefore, we set our objective as virtually reducing the issue of optical scaling

effect, and we set our technological issue as the construction of a “real-time all-in-focus microscope,” which is a new micro-visual system that solves the optical scaling effect within “real time” during microscopic manipulations in high-power optical systems such as the optical microscope^{[1]-[4]}.

2 History of struggle: twists and turns

Here, we chronologically present the path from an idea that initially started as a scientific interest to product realization, while overcoming several feasibility study (FS) phases. The primary innovative idea was published in regular academic papers and as product articles after it was finalized as a product. For the *Synthesiology* paper, I shall explain the process from the idea phase through several FS phases and then finally to product realization. I shall specifically mention that the course of these FS phases was greatly affected by the “strategic deepening and selective synthesis” that are described as synthesiology including selections and rejections of some components, development of algorithms dedicated to those components, and meeting by chance with partners to realize them, as well as the “lucky coincidences” that couldn’t be quite categorized. That is because as we blindly wandered in the dark and finally arrived at the goal,

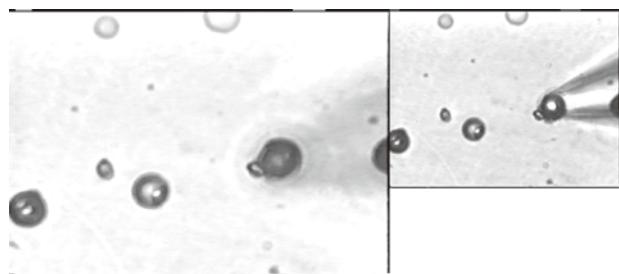


Fig. 1 Microscope image of microscopic operation (manipulating 4 μm glass beads).

(a) Glass beads and forceps are at different heights, and (b) glass beads and forceps are at same height.

Intelligent System Institute, AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba 305-8568, Japan E-mail : k.ohba@aist.go.jp

Original manuscript received December 5, 2008, Revisions received August 20, 2009, Accepted August 21, 2009

only in retrospect can we say certain places were the forks in the road. While I was groping desperately, honestly speaking, I cannot say I did any strategic decision-making even in retrospect. Even if a theoretical order was given for the twists and turns, the situation differed greatly in the next blind wandering, and there were very few instances when I felt the previous experiences helped at all. Therefore, our history of twists and turns (Fig. 2) will be presented as an article that documents a process of product realization, not in the format of an academic thesis, and I shall explain the phases from a technological standpoint in the following sections.

(a) Idea phase

When a person looks at an object with the naked eye, the objects both close and afar seem to be in focus. This is accomplished by the focal adjustment of the eyes. In contrast, when one sees things through the lens such as of a camera, it is necessary to bring the lens into focus. The automatic focus camera sets the focus automatically, but it can do so at certain distance only.

In a microscopic environment, the depth of field can be increased by narrowing the aperture in the low-power single-lens reflex cameras and stereomicroscopes, but it is impossible to cover all areas as magnification increases due to the optical principle. In microscopic objects, the depth

of the objects cannot be discerned through images with deep depth of field. Therefore, we thought of increasing the operability in the microenvironment by making use of the shallow depth of field.

To increase the operability in the microenvironment, we decided to create a system that fulfills the following two conditions simultaneously, and conducted theoretical considerations and devised processing algorithms:

1. Dynamic observation (30 frames/sec.) of real image with depth of field ideally raised to infinity
2. Real time measurement (30 frames/sec.) of three-dimensional shape of the object

At this time, we were vaguely thinking that it may be interesting to have this function in the digital camera or eyeglasses, and had no idea about any specific device application.

The above conditions 1) and 2) were not stated as initial target specifications, but as we were looking through papers for methods to obtain the image of 1) using the image processing technology, we realized that the 3D position information of objects were not used in the course of processing. Therefore, we thought why not use this information? The details will be discussed later.

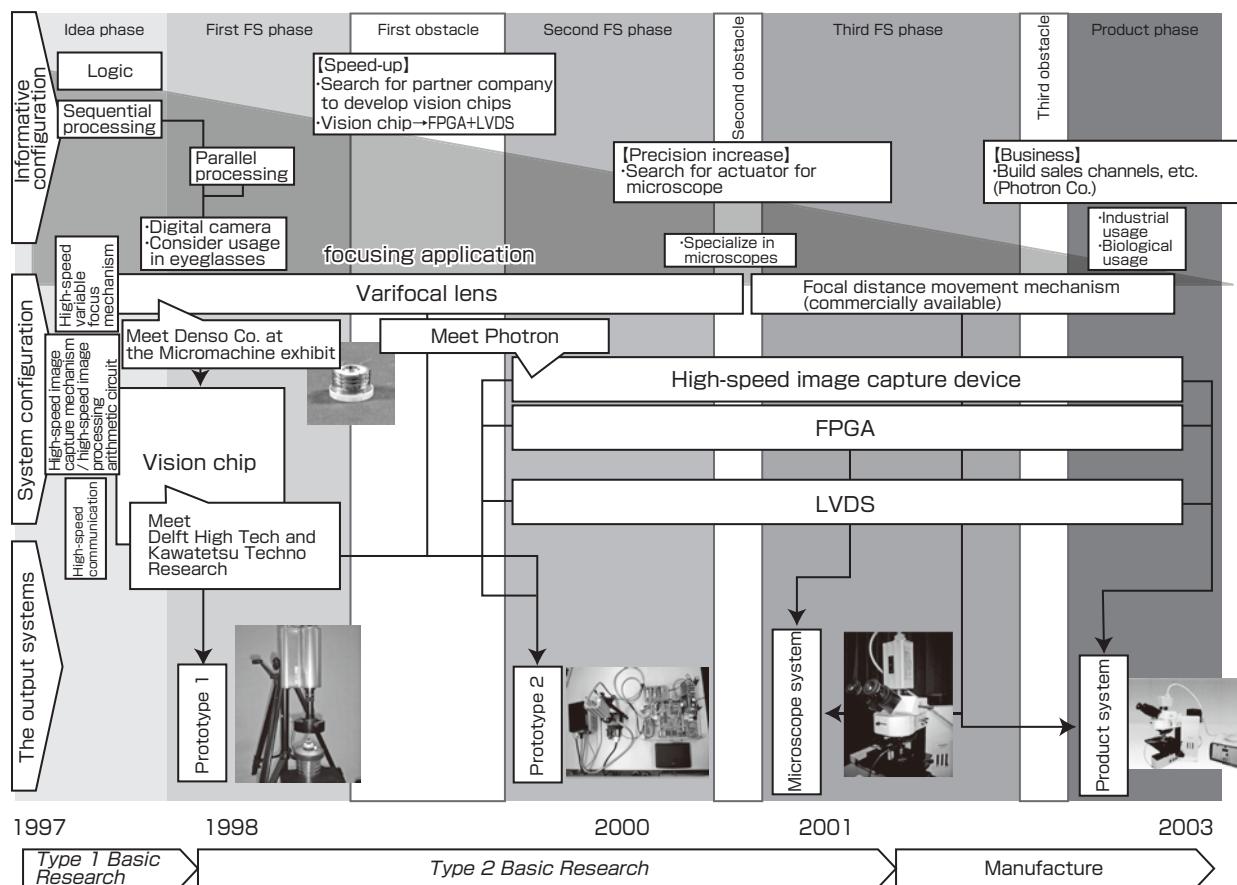


Fig. 2 Our struggle (Twists and turns) for synthesis.

(b) First FS phase

To incorporate the idea into an actual hardware system, high-speed variable focus mechanism and information processing and communication technology to accomplish the data acquisition, communication, and processing in real time were required to synthesize the system. For the high-speed variable focus mechanism that allowed dynamic observation at 30 frames/sec., response rate at 30 Hz was required, but available products only had single-digit response. However, this was solved when we visited a micromachine exhibition and came across the varifocal lens developed by the Denso Corporation in their micromachine project. For information processing and communication, when the number of necessary images to composite a single image was N , $N \times 30$ frames/sec. \times image digital size would be necessary. We then met Delft High Tech Corporation (current DHT Corporation) that imported the vision chip MAPP series from the Netherlands, and Kawatetsu Techno Research Corporation (current JFE Techno Research Corporation) that was working on its development. We were able to solve both technological issues and created Prototype 1.

(c) Second FS phase

However, speed remained at 0.5 frames per sec., and the initial goal of 30 frames/sec. could not be achieved. We spent about a year negotiating with several companies to develop the high-speed vision chip. This was our first obstacle. In ordinary research, we could have written, "The problem can be solved by developing an ultra high-speed vision chip and implementing it to the device," and then leave it. Yet, we set out on the voyage to clear the first obstacle from the desire to carry this project out to the end. However, even if we visited the companies with our request, they turned us down because they did not see any business merits for spending several hundred million yen for the development. The break came when we visited Photron Ltd. This was a high-speed camera manufacturer that had technology to shoot images at 10,000 frames/sec. and to transfer them to memory. They gave us encouraging advice that our requirement could be realized by a high-speed image capture device and the LVDS (low voltage differential signaling)^{Term 1} interface and by using the FPGA (field programmable gate array)^{Term 2} without developing the vision chip. Prototype 2 was created in a month.

(d) Third FS phase

In this phase, we started discussing future business with Photron. When we participated in exhibitions, we learned that the all-in-focus image was most highly demanded in microscope use, and therefore we started developments specializing in microscope application. Initially, the Denso varifocal lens was hand-made, mass production was not possible, and it was not suitable for high-precision positioning required for the microscope. We sought out the piezo-actuator of Physik Instrumente (PI) GmbH of Germany. In this device, the focal distance of the microscope was adjusted

electrically, and we could obtain the desired precision by modifying the control unit. The second obstacle was to achieve high precision in high-speed mechanical movement, and the third FS phase was to overcome this issue. Initially, we were thinking only about the high-speed information processing, and expected the high-speed mechanical movement would somehow fall into place. In fact, we were plagued to the end on how to achieve highly reliable high-speed mechanical movement.

(e) Product realization phase

By repeating several demonstrations using the microscope system, we considered further business application by separating the uses into industrial and biological. However, the microscope system had a special sales channel where the optical manufacturer visited the researchers to sell it and then became in charge of maintenance. For an outsider like Photron to enter, it had to provide OEM to conventional optical companies and establish the sales channel to reach the users. This was our final obstacle. Over one year was needed to establish the sales channel, and the product was launched into the world in 2003.

3 Idea phase

The idea explained in the previous section was born from a simple scientific interest of "wanting to make an image with everything in focus." The initial application of this function was vaguely considered for digital cameras and eyeglasses. We started by checking the efficacy by conducting offline processing on a PC for several images with different focal distances.

In this phase, to conduct simple processing within the memory capacity and processing ability of the PC, we developed a sequential processing algorithm to increase the memory efficiency and a parallel processing algorithm to increase the processing capacity.

3.1 Theoretical synthesis - depth from focus method

It was described earlier that in case of optical microscope image, the issue of shallow depth of focus greatly affected the operability. At the same time, this may be a major advantage to realize the depth from the focus method^{[6][7]} which is one of the 3D measurement methods of objects^[5]. While the all-in-focus camera was a solution to the shallow depth of focus, by making use of this problem, not only could we achieve an all-in-focus image where every part was in focus, but also by using the depth from the focus method, it became possible to obtain the 3D shape of the object using only a single lens.

Figure 3 shows the conceptual diagram of the depth from the focus method. When observing an object with different depths, the distance to the image surface is scanned, and the local frequencies of the dark-light data at each point of the

image are measured to detect the peak. Once the distance to the image surface in focus is obtained, the distance to the object can be calculated by using the Gauss' lens law.

Whether the focus is correct can be determined by conducting the local spatial frequency analysis around the observation point of the image while moving the focal distance f , object distance l , or image distance l' . The point with greatest spatial frequency is the point in focus. This method is used often in the automatic focus mechanism, and one can intuitively see that the out-of-focus area has low frequency while the in-focus area has high frequency. Basically, the focus of the lens is moved using the variable focus mechanism. The images are captured one by one, the local spatial frequency analysis around the pixel point is conducted for each image, the peak of the frequency or the area in focus are picked up from the images for each pixel, and these are pasted together as one image to obtain an all-in-focus image. The 3D data can also be obtained from the focal distance and image distance at each point.

There are many methods of assessing the degree of focus of an image such as looking at the changes in brightness of the image while changing the focal distance. In this paper, considering the final product realization, we define the following equation for Image Quality Measure (IQM), to assess the local spatial frequency analysis of each pixel through spatial dispersion of image brightness value, for the reason that the image-processing algorithm can be easily implemented in the hardware. This IQM value was originally defined as one of the indices that indicate the clarity of the image, and was not for determining whether the image is in focus or not. However, we decided to use the IQM value since the processing algorithm can be easily adapted to higher speed in the future, with the assumption that the image is digitized and will be digitally processed.

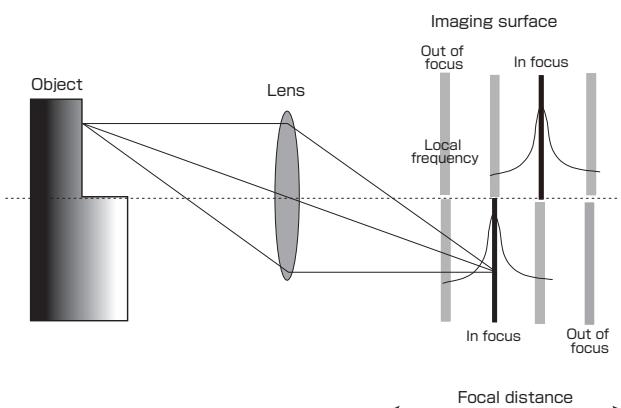


Fig. 3 Depth from focus method.

$$IQM = \frac{1}{D} \sum_{x=x_i}^{x_f} \sum_{y=y_i}^{y_f} \left\{ \sum_{p=-L_c}^{L_c} \sum_{q=-L_r}^{L_r} |I(x, y) - I(x+p, y+q)| \right\}$$

Here, $(-L_c, -L_r) - (L_c, L_r)$ and $(x_i, y_i) - (x_f, y_f)$ are small regions for conducting dispersion assessment and smoothing. D is the number of all pixels to which assessment is conducted for normalization at pixel unit. The IQM values are assessed for each pixel or region while moving the focal distance, the peak of the IQM value is detected, the object distance l is calculated from the focal distance f and image distance l' , and then this is substituted in the matrix component for each pixel position to create the 3D data of the object.

3.2 Configuration of sequential processing – for reduced load on memory

As mentioned before, it is theoretically possible to obtain both the all-in-focus image and the depth image simultaneously, using the depth from focus method. However, in 2000 when we started this development, when calculating the algorithm for the IQM value, 2 Mbyte image memory and capture and processing of 30 shots/sec. \times 30 frames = 900 images (about 3 min. using the PC in 2000) were necessary to obtain one all-in-focus image and depth image from about 30 images of 256×256 pixels in real time.

In case it is necessary to obtain N number of different depth images to capture an all-in-focus image and a depth image at 30 frames/sec., an image capture device with high dynamic range that can shoot at $30 \times N$ frames/sec. is required. Moreover, a high-speed processing system to process and display such volume of image data is required.

In an automatic focus camera, to obtain the IQM value, the values for one or few points can be calculated and the focus can be moved according to the value. However, to obtain the all-in-focus image, calculations must be done efficiently within 33 ms for each pixel point.

The method the authors devised for optimizing the memory constitutively through algorithm and for overcoming the limitation of this hardware will be described below. In the following chapter, the configuration for optimizing the processing speed using the hardware characteristics will be explained.

When conducting these IQM processing at all pixel points, it was not efficient to process by temporarily storing the different pre-images with varying focal distances. Therefore we constructed a configuration using the sequential algorithm with steps (1)-(7) as shown below. Figures 4 and 5 show the main system configuration diagram and the flow chart, respectively.

Using this sequential algorithm, sequential processing is conducted while changing the focal distances, and by moving the focal distance to the end, each of the matrixes of the finally updated image memory becomes the all-in-focus image (AIF) and the depth image (DEPTH). While the equation for IQM seems complicated, only Laplacian and smoothing processes are applied as image processing technology. Laplacian is a secondary differentiation, and in the world of digital image, differentiation is the difference from the neighboring pixel, and it becomes Laplacian by differentiating twice. Smoothing is averaging. The two processes are four arithmetic operations, and are optimal for hard logic circuit for high speed. Normally, the memory will hold N frames of differing depth images, the IQM image for each are calculated, and the all-in-focus and depth images are obtained (image memory for total $2N + 2$ frames were necessary) by comparing the IQMs at the same pixel position. In contrast, in the sequential algorithm, the obtained images are sequentially compared on the spot, and the memory is needed only for four images consisting of the original image, IQM image, all-in-focus image, and depth image. This reduced the memory requirement, and it became easy to configure the hardware, as it will be explained later.

- (1) Format
init (IQM);
- (2) Move focal distance
for (FV=0 to FVMAX) { (FV: focal distance)
 mov (FV); move focal distance
- (3) Image input
 IO = input; image input
- (4) Laplacian filter
 IL = lap (IO); preliminary processing
- (5) Median filter (average filter)
 IM = ave (IL); preliminary processing
- (6) Comparison of brightness value of each pixel, and corresponding copy of image data
 if (IM(x, y) > IQM(x, y)) {
 IQM(x, y) = IM(x, y); update assessment value
 AIF(x, y) = IO(x, y); create all-in-focus image
 DEPTH(x, y)=FV; create focal distance information
 }
}
- (7) Image data output
 output (AIF, DEPTH); image data output

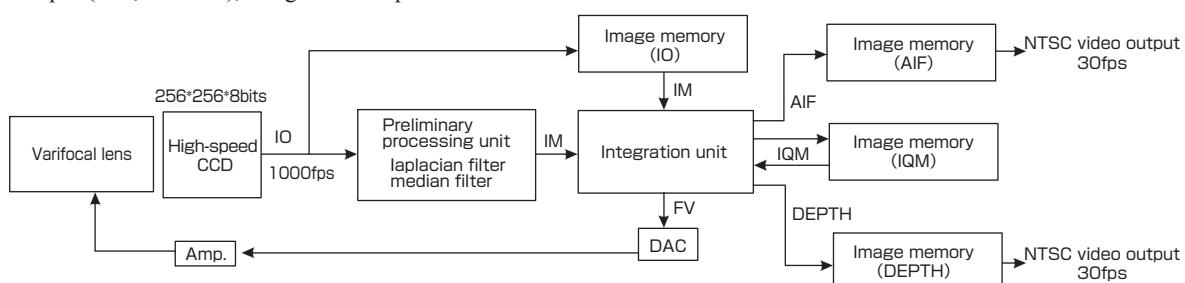


Fig. 4 System configuration for sequential algorithm.

4 First FS phase

To accomplish the above methods and information processing in real time, it is necessary to move the depth of focus at high speed and at the same time, capture and process the images at high speed. Therefore, we determined that the basic system components were the following three:

- (a) high-speed variable focus mechanism
- (b) high-speed image capture mechanism, high-speed image processing arithmetic circuit
- (c) high-speed communication.

For example, in conducting the process in eight-step focal distance, to obtain the output in real time at 30 frames/

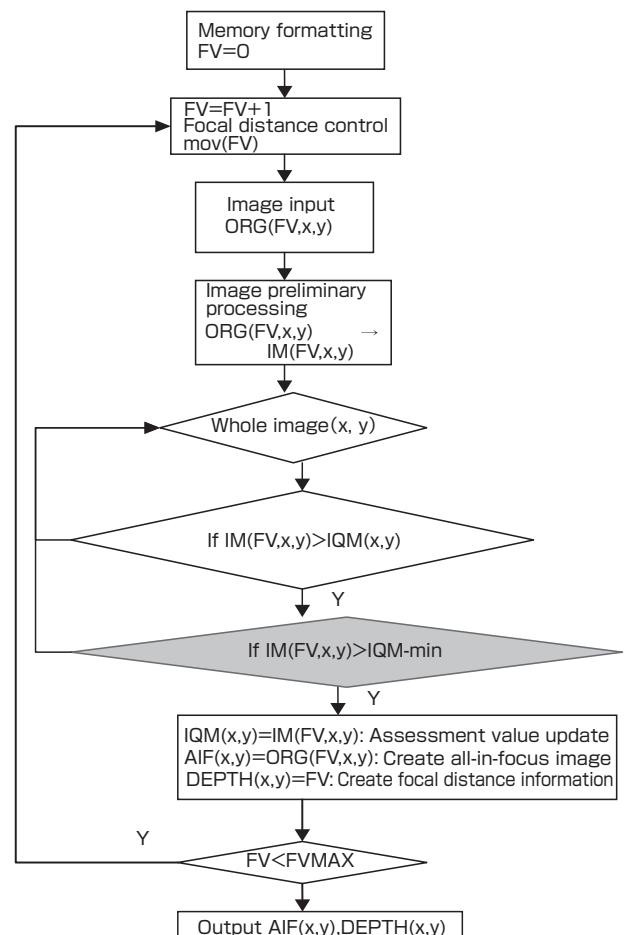


Fig. 5 Flowchart for sequential algorithm.

sec. that is sufficiently smooth for human observation, synchronized high-speed movement of the focal distance at 30 Hz and the image capture and processing speed of 240 frames/sec. at 30×8 are required (Fig. 6). In case of 240 frames of black-and-white 512×512 pixels, the pixel rate will be close to 100 MHz. At the same time, the focal distance of camera, the object distance, or the image distance must be moved at 30 Hz.

(a) High-speed variable focus mechanism

For Prototype 1 and Prototype 2, the varifocal lens developed by Denso Corporation was used as the high-speed variable focus mechanism^[8]. It is driven by a piezo element, and the focal distance changes according to the voltage applied. The structure is simple without any motor. The focal distance is changed by moving the glass diaphragm using the bimorph actuator. By changing the voltage applied to the PZT bimorph, the lens can be changed from a convex to a concave lens. It has been demonstrated that the frequency response is possible up to around 150 Hz without delay in phase. Figure 7 shows the varifocal lens, Fig. 8 shows the structure, and Fig. 9 shows the details of the lens driving mechanism. When no voltage is applied, the lens is a planar glass. The greatest characteristic of this varifocal lens is its high speed. Since the glass diaphragm is directly driven using the piezo element, high-speed movement of the focal distance is possible.

(b) High-speed image capture mechanism, high-speed image processing arithmetic circuit

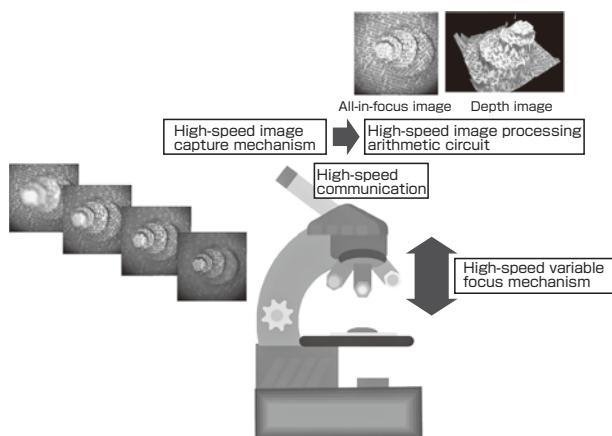


Fig. 6 System configuration for all-in-focus microscope.



Fig. 7 External appearance of varifocal lens.

For the high-speed image capture mechanism and the high-speed image processing arithmetic circuit in Prototype 1 of section 3.1, we used the vision chip that included an image capture device, ADC, and a processing system. By using this vision chip, processing was concluded within the vision chip, so high-speed communication was unnecessary. Here, since the volume of image data per unit time and image processing capacity were high, we used the column parallel type vision chip, MAPP2200, from Integrated Vision Products AB of Sweden. The basic configuration was CMOS image sensor with 256×256 pixels, 256 ADCs, and 256 parallel processors. As mentioned in the previous chapter, using the image-processing algorithm that allows SIMD (single instruction multiple data)^{Term 3} processing, the image capture and processing can be done at high speed.

(c) Configuration with parallel processing

In 2000, when we started the development, although the image-processing technology at an ordinary frame rate was already in practical use, the image capture and processing at a frame rate one digit higher would not be realized unless a special vision chip was developed. Of course, we were told that development of such vision chip required about a hundred million yen in cost. We visited several companies, but none gave us encouragement. In Prototype 1, we used the general-use vision chip MAPP2200 that was developed in Sweden to implement the parallel processing algorithm. Two-second processing speed was realized, but this was far from the real-time processing at our target frame rate.

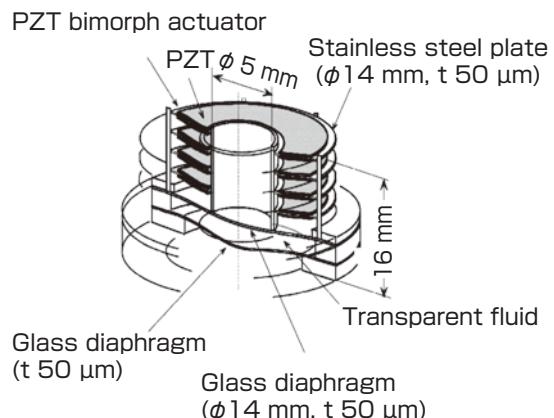


Fig. 8 Configuration of varifocal lens.

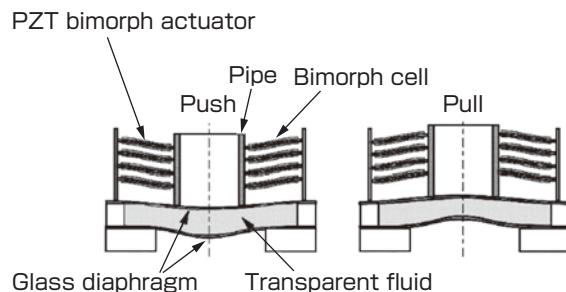


Fig. 9 Principle of varifocal lens function.

High processing capacity was required for the Laplacian filter and smoothing filters in the preliminary processing of the image in this method. Figures 10 and 11 show the hardware version of this process, which is an example of the vision chip implemented in Prototype 1 that will be explained later in this chapter. For the hardware used, the chip allows easy SIMD processing, and therefore, to calculate the above IQM value, as a process using the pixel values around each pixel points, parallel calculation of the Laplacian and smoothing filters are necessary. The examples of SIMD processing for each operation are shown. Looking at each calculation result, Laplacian filter (R12 of Fig. 10) and smoothing (R5 of Fig. 11) were ultimately obtained for each pixel point.

For the final product realization, while using this processing method as a reference, we configured a system that calculated the IQM value using the FPGA. The external appearance of Prototype 1 system created in 1999 is shown in Fig. 12.

As an assessment test, we conducted processing by placing an object with a depth of 35 mm at a position of 160 mm from the varifocal lens. To move the focal distance to cover an object with a height of 35 mm, 21 images were captured, and the resolution of depth was 1.67 mm. While the spatial resolution of the system is dependent on the setting of the optical device, in this case, since $16 \text{ mm} \times 16 \text{ mm}$ was processed at an image resolution of 256×256 , the resolution was $62.5 \mu\text{m}/\text{pixel}$.

	0	255	
R1	$I(i-1, j-1)$	$I(i-1, j)$	$I(i-1, j+1)$
R2	$I(i, j-1)$	$I(i, j)$	$I(i, j+1)$
R3	$I(i+1, j-1)$	$I(i+1, j)$	$I(i+1, j+1)$
R4=R2-R1	$I(i, j-1)-I(i-1, j-1)$	$I(i, j)-I(i-1, j)$	$I(i, j+1)-I(i-1, j+1)$
R5=R3-R2	$I(i+1, j-1)-I(i, j-1)$	$I(i+1, j)-I(i, j)$	$I(i+1, j+1)-I(i, j+1)$
R6=R5-R4	$\frac{I(i+1, j-1)+I(i-1, j-1)}{-2I(i, j-1)}$	$\frac{I(i+1, j)+I(i-1, j)-2I(i, j)}{2I(i, j)}$	$\frac{I(i+1, j+1)+I(i-1, j+1)}{-2I(i, j+1)}$
R7=ROR(R2)	$I(i, j-2)$	$I(i, j-1)$	$I(i, j)$
R8=LOR(R2)	$I(i, j)$	$I(i, j+1)$	$I(i, j+2)$
R9=R7-R2	$I(i, j-2)-I(i, j-1)$	$I(i, j-1)-I(i, j)$	$I(i, j)-I(i, j+1)$
R10=R8-R2	$I(i, j)-I(i, j-1)$	$I(i, j+1)-I(i, j)$	$I(i, j+2)-I(i, j+1)$
R11=R9+R10	$I(i, j)+I(i, j-2)-2I(i, j-1)$	$I(i, j+1)+I(i, j-1)-2I(i, j)$	$I(i, j+2)+I(i, j)-2I(i, j+1)$
R12=R6+R11	$I(i+1, j-1)+I(i-1, j-1) + I(i, j)+I(i, j-2)-4I(i, j-1)$	$I(i-1, j)+I(i+1, j) + I(i, j+1)+I(i, j-1)-4I(i, j)$	$I(i+1, j+1)+I(i-1, j+1) + I(i, j+2)+I(i, j)-4I(i, j+1)$

Fig. 10 Laplacian filter processing operation by SIMD processing.

The object used here was an artificial four-step pyramid of a height of 35 mm. The first step was $\varphi 10 \text{ mm}$ and a height of 10mm, the second step was $\varphi 7 \text{ mm}$ and a height of 10 mm, the third step was $\varphi 4 \text{ mm}$ and a height of 10 mm, and the fourth step was $\varphi 3 \text{ mm}$ and a height of 5 mm.

Part of the 21 images shot while moving the focal distance is shown in Fig. 13 on the left. As mentioned earlier, by conducting processing, the all-in-focus image as in Fig. 13 on the right and the VR display as in Fig. 14 were obtained. While the all-in-focus image itself was adequate, we obtained a largely dispersed result since the depth image had a small region that required smoothing, and the resolution in the depth direction was small.

While the performance of Prototype 1 was dependent on the

R1	a b c d e f g h i j k l
R2=cros(R1, 1)	b a d c f e h g j i l k
R3=R1+R2	a+b a+b c+d c+d e+f e+f g+h g+h i+j k+l k+l
R4=cros(R3, 2)	c+d c+d a+b a+b g+h g+h e+f e+f k+l k+l i+j i+j
R5=R3+R4	a+b+c+d e+f+g+h k+l+i+j

cros(R1, 1)	R1 a b c d e f g h i j k l
R2=ROR(R1)	a b c d e f g h i j k
MASK(1)	0 1 0 1 0 1 0 1 0 1 0 1 0 1
R3=R2&MASK(1,R)	0 a 0 c 0 e 0 g 0 i 0 k
R4=ROL(R1)	b c d e f g h i j k l a
MASK(1)	1 0 1 0 1 0 1 0 1 0 1 0 1 0
R5=R4&MASK(1)	b 0 d 0 f 0 h 0 j 0 1 0
R6=R3 or R5	b a d c f e h g j i l k

MASK(1)	0 1 0 1 0 1 0 1 0 1 0 1 0 1
MASK(2)	0 0 1 1 0 0 1 1 0 0 0 1 1 1
MASK(3)	0 0 0 0 1 1 1 1 0 0 0 0 0 0
⋮	⋮

Fig. 11 Summation processing operation by SIMD processing.



Fig. 12 Prototype 1.

setting of the area size and the number of captured images, output of about one image per 2 sec. could only be obtained as the processing time. Since the MAPP2200 had ability to capture and process binary images at 2,000 to 3,000 images per second, the reasons for the slow processing speed were thought to be: 1) it was necessary to conduct sequential comparison by providing a reference voltage 256 times when capturing an image at 8 bit resolution, but the column A/D could not provide individual reference voltage for each pixel, and 2) the architecture of the SIMD processor was specialized for binary images.

5 Second FS phase

Here, I shall explain the second FS phase for realizing Prototype 2 of the all-in-focus camera that allows real-time observation at 30 frames/sec..

In the microscope system, we were able to obtain cooperation of Photron Ltd., a company with abundant experience in high-speed image capture. In this system, the camerahead of the high-speed video camera was used, and by using LVDS that allowed high-speed image transfer as the interface between the imaging mechanism and the image processing arithmetic circuit, the imaging mechanism and the image processing arithmetic circuit were separated. The commercially available high-speed imaging mechanism (high-speed video camera) and the image processing arithmetic circuit (FPGA) were used. It can be said that the

road to product realization opened up widely by using the existing products.

The configuration of Prototype 2 is shown in Fig. 15 and the external appearance is shown in Fig. 16. The output of the high-speed sensor goes through the CDS (correlated double sampling) and ADC, converted by the high-speed digital interface LVDS, and then transferred to the image-processing unit. Up to this point is the description of the part of the high-speed camera. LVDS is a standard interface for sending high rate image signals, and is a standard often used in digital LC displays. The saw-tooth generating circuit for the varifocal lens receives the synchronizing signals from the clock generator of the high-speed camera part, produces saw-tooth pulses at 30 Hz, and drives the lens with the lens-driving amp. In the image processing part, the IQM calculation and image compositing are accomplished with the input digital image signals, and outputted as VGA (video graphics array) signals. The 3D data is transferred to the PC by LVDS signals. At the PC, data is received by the LVDS capture board of the PCI bus.

Mechanically, the focal distance movement used was the same as the one used in Prototype 1 and there was no problem. However, since the algorithm was implemented on the special vision chip in the first FS phase, it could not be transplanted directly to the FPGA, and it was necessary to modify the algorithm for FPGA. It was also necessary to use the internal memory to speed up the FPGA processing, and

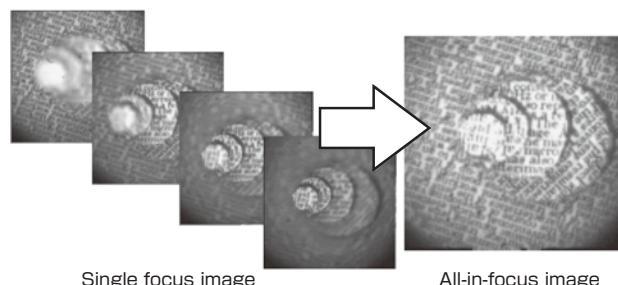


Fig. 13 Schematic diagram for the creation of all-in-focus image.

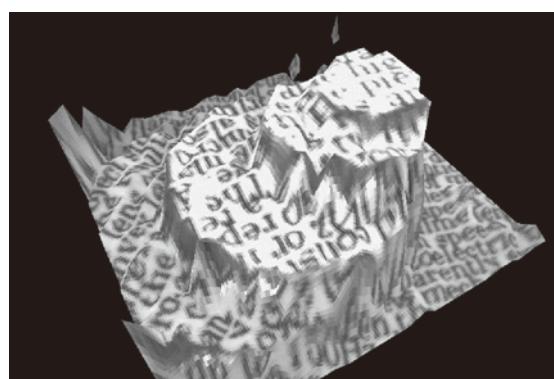


Fig. 14 Example of VR display.

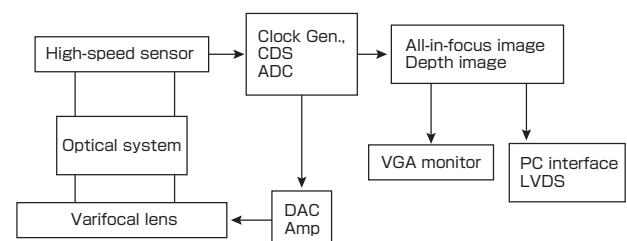


Fig. 15 Block diagram for Prototype 2.

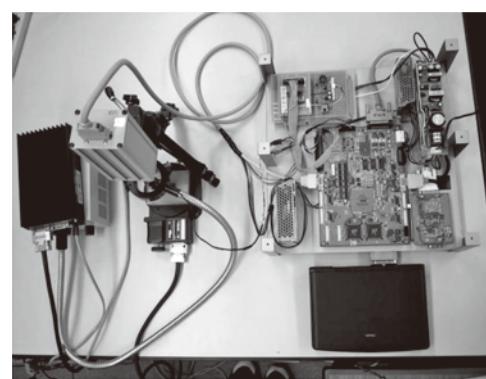


Fig. 16 External appearance of Prototype 2.

implementation to increase the memory capacity became necessary. Here, implementation of the sequential processing described earlier reduced the memory volume. As a result, processing could be accomplished within the internal memory of the FPGA, and target specification of 30 frames/sec. was achieved.

Figure 17 shows the examples of output images using Prototype 2. They look the same as the ones shown for Prototype 1, but actually, the all-in-focus images shown on the left and the depth images shown on the right are moving images. Eight focal points were recorded per frame, and as the performance of Prototype 1 was 0.5 frame/sec. while that of Prototype 2 was 30 frames/sec., the processing speed of the latter was 60 times faster.

6 Third FS phase

With the developments up to this point, the following points became apparent.

- By using the existing camerahead of high-speed cameras and LVDS, image capture, communication, and processing could be separated, and high-speed parallel processing using general FPGA was possible for the processing part.
- Mechanical focal distance movement could be done using varifocal lens.

From the talks with the company, while utilizing the high-speed processing part, we decided to specialize in the microscope use that had the highest potential in terms of business application. However, for the mechanical focal distance movement, the second item mentioned above, the varifocal lens could not be used due to precision issues. Here, I shall explain the development of the system specifically for microscope use in the third FS phase.

For the mechanism for moving the focal distance in the microscope system, we used the commercially available focal

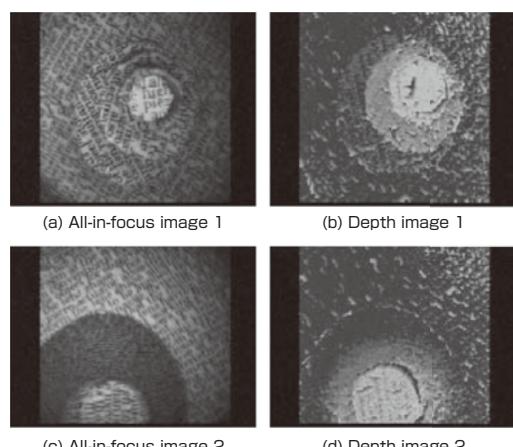


Fig. 17 Example of output by Prototype 2.

distance movement mechanism PZT actuator P-721 and 20 and driver E-612, C0 from PI Polytec Inc. By attaching these between the objective lens of the microscope, they enable parallel movement of the objective lens for 0-100 µm. The reason for not using the varifocal lens used in the prototypes of chapters 4 and 5 was because in the varifocal lens, the *f* value of the lens itself was changed by changing the thickness of the lens. Therefore strictly speaking, images of different magnifications were composited, and this was not appropriate for producing high precision all-in-focus images.



Fig. 18 External appearance of microscope system.

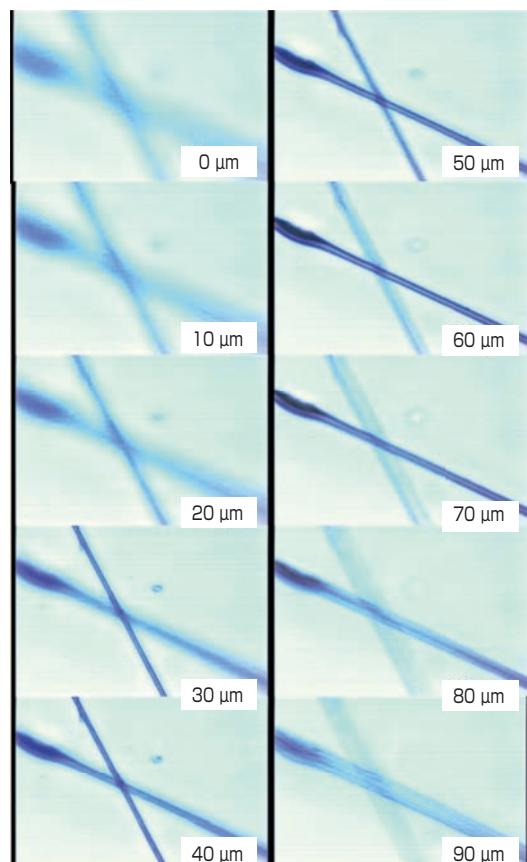


Fig. 19 Example of typical microscope image.

Figure 18 is an external appearance of the microscope system implemented on the microscope using Prototype 2 as a base. Figure 19 shows examples of ordinary optical microscope images, and Fig. 20 is an image obtained with the microscope system.

Here, glass fibers of 2 μm diameter that are crossing each other three dimensionally are used as measured objects. Figure 19 shows the image every 10 μm . It can be seen that the first glass fiber is in focus at 30-40 μm while the second glass fiber is in focus at 60 μm .

By using the proposed all-in-focus microscope camera, an all-in-focus microscope image, shown in Fig. 20, was obtained. Since this image is a movie, the all-in-focus image is sequentially updated even if the measured objects move. The depth image is also obtained.

7 Product phase

The system developed here was realized as a product by Photron in fiscal year 2003. The external appearance of this system is shown in Fig. 21. The basic configuration does not differ from the prototype, but measures are taken for practical use such as the incorporation of color images and keeping the size of the system as small as possible.

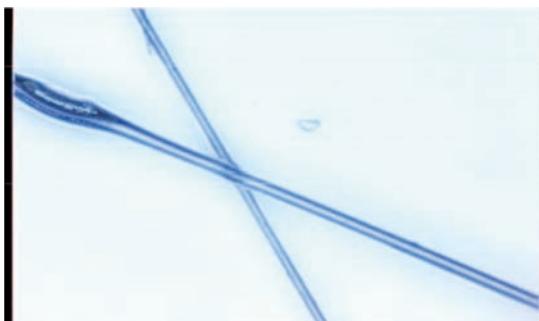


Fig. 20 Example of all-in-focus microscope image (moving image).

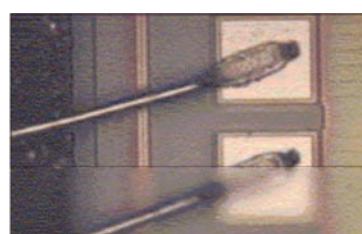


Fig. 21 External appearance of product system.

The examples of all-in-focus images obtained with this system are shown in Fig. 22, in comparison with ordinary images. Figure 22 top shows the wire bonding of an IC chip, and Fig. 22 bottom shows the diatom and microbes by transmitted light. For both images, the objective lens was of $\times 50$ magnification. In actual performance, up to about $\times 100$ objective lens can be operated, with the limiting factor being the weight of the objective lens. The stereomicroscope is used for low-power magnification from about 1 mm to 0.1 mm, and the electron microscope is widely used for high-power submicron observation. Due to the optical limit of the optical microscope, high-power is not needed in the biological field or for semiconductor inspection. Therefore, it was confirmed that the above magnification was sufficient for practical use.

8 Future issues

The all-in-focus microscope camera solves the issue of shallow depth of field that is a problem in ordinary microscope images, as well as the issue of not being able to obtain depth information. It is a system whose objective is to produce and display images where all areas are in focus and to conduct 3D compositions of objects in real time. By allowing all-in-focus images and 3D compositions at the same time, an object can be observed in detail. However, by specializing in extracting only the image in focus and the depth data from one direction only among the multiple images of varying focal distances obtained by ordinary microscope, much information had to be cut off. For example, when the focal distance is moved in a vertical direction for a



All-in-focus image



Ordinary image



All-in-focus image

Ordinary image

Fig. 22 Example of image output of product.

translucent object, while there can be an object with several points in focus in the vertical direction, this method allows selection of one point only and unconditional observation of other areas become impossible.

As another method for reducing the optical scaling effect, the technology for volume rendering^{Term 4} directly from the images with multiple focal distances, as shown in Fig. 23, was considered. This has been patented^[9].

In the future, it is expected that there will be demand for a hyper-microscope that can produce image slices from any direction by volume rendering all the depth images, without having to consider the hardware limitations.

Acknowledgements

This study was based on the joint research by the Mechanical Engineering Laboratory, Agency of Industrial Science and Technology (current AIST), Ministry of International Trade and Industry; Delft High Tech Corporation (current DHT Corporation); Kawatetsu Techno Research Corporation (current JFE Techno Research Corporation); and Denso Corporation. The joint research for practical application was conducted by AIST and Photron Ltd. Part of the varifocal lens used in this study was manufactured by Denso Corporation, as a subcontract of the Micromachine Center that was subcontracted by the New Energy and Industrial Technology Development Organization (NEDO), as part of the “R&D for Micromachine Technology” based on the Industrial Science and Technology Frontier Program, Agency of Industrial Science and Technology, Ministry of International Trade and Industry.

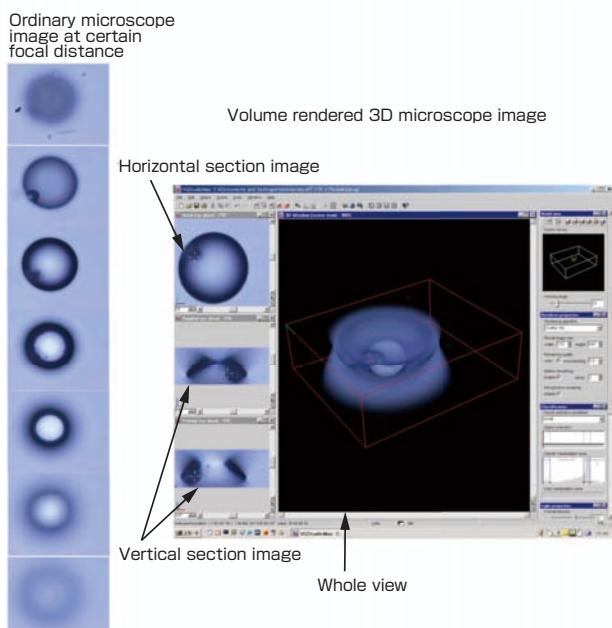


Fig. 23 Overview of volume rendering method.

Terminology

- Term 1. LVDS (low voltage differential signaling) interface: standard of electric signal that enables high-speed function using twist pair cables. It has been used in high-speed cameras to handle high volume data. It is used in PCs as interface with LC displays.
- Term 2. FPGA (field programmable gate array) processing: a type of programmable gate array where the users can write their original logic circuit. The gate array is arranged in two-dimensional lattice form. It is suitable for parallel processing operation.
- Term 3. SIMD (single instruction multiple data) processing: process in an arithmetic device where simultaneous processing for several data is done with one command.
- Term 4. Volume rendering: When expressing a 3D object as a 2D image, rather than the method where the depth is expressed by adding shadow on the object surface, in this method, the 3D object is made to look three dimensional by adding transmittance and color information inside the object.

References

- [1] K. Ohba, J.C.P. Ortega, K. Tanie, G. Rin, R. Dangi, Y. Takei, T. Kaneko and N. Kawahara: Implementation of real-time micro VR camera, *Denki Gakkai Ronbunshi E (IEEJ Transactions on Sensors and Micromachines)*, 120-E (6), 264-271 (2000) (in Japanese).
- [2] K. Ohba, J.C.P. Ortega, K. Tanie, G. Rin, R. Dangi, Y. Takei, T. Kaneko and N. Kawahara: Micro-observation technique for tele-micro-operation, *Advanced Robotics*, 15 (8), 781-789 (2001).
- [3] K. Ohba: Frame-rate all-in-focus microscopic system, *Nihon Robotto Gakkaishi (Journal of Robotics Society of Japan)*, 21 (1), 43-44 (2003) (in Japanese).
- [4] Patent 3737483, Real-time all-in-focus microscope camera
- [5] M. Ishihara and T. Furusawa: Saikin no hikari hyomen keijo keisoku gijutsu, *O plus E*, 20 (11), 1251-1258 (1998).
- [6] S.K. Nayer and Y. Nakagawa: Shape from focus, *IEEE Trans. on PAMI*, 16 (8), 824-831 (1994).
- [7] K. Kodama, T. Onishi, K. Aisawa and M. Hattori: Arbitrarily focused image acquisition by using multiple differently focused images, *the Journal of the Institute of Image Information and Television Engineers*, 51 (12), 2072-2081 (1997).
- [8] T. Kaneko, N. Taya, N. Kawahara, N. Akita and T. Hattori: Kahlen shoten lenzu o mochiita cho shoten shindo shikaku kiko, *Denki Gakkai Maikuro Mashin Kenkyukai* (1997).
- [9] Patent 3627020, Three dimensional transmissive microscope system and image display method

Author

Kohtaro Ohba

Completed doctorate at the Graduate School, Tohoku University in 1991. Doctor (Engineering). Deputy director of Intelligent Systems Research Institute, AIST and group leader of Dependable System Research Group from 2009. Currently studying ubiquitous robot and dependable system in pursuit of robots that can be actually used in everyday living. Member of the Institute of Electrical and Electronics Engineers (IEEE) and the Japan Society for Mechanical Engineers. Also teaches as Professor, Cooperative Graduate School, University of Tsukuba; Professor, Cooperative Graduate School, Shibaura Institute of Technology; and Visiting Associate Professor, Graduate School of Interdisciplinary Information Studies, the University of Tokyo.



Discussion with Reviewers

1 Synthesis method

Question and comment (Naoto Kobayashi, Center for Research Strategy, Waseda University)

In this paper, the author has conducted an information configuration based on the original idea of the “configuration of real-time all-in-focus image” in the first part, then has conducted a system configuration including the hardware to realize the “configuration of real-time all-in-focus image,” and ultimately has achieved a product realization of the real-time all-in-focus microscope by clearly improving the objectives sequentially. This is extremely important. Here, we see that the extremely unique R&D unseen elsewhere, despite its twists and turns, has been conducted and completed according to the following steps: 1) consistency from *Type 1 Basic Research* to product realization, 2) sequential deepening and clarification of the strategic goals, 3) improvement of the required system configuration, and 4) actual product realization and maintenance as a commercial product thereafter. I thought these were appropriate as a *Synthesiology* paper.

On the other hand, in *Synthesiology*, the originality of the synthesis method is one of the key issues. For this paper, I guess that the strategy became clearer as the technology or the synthesis progressed, and the scenario became more apparent as it proceeded. I understand that there was no clear strategic objective such as “development and product realization of real-time all-in-focus microscope” in the beginning, but it fell into a cycle where the strategic goal deepened as the elemental technologies were synthesized, the issues to be solved in the next step became clear, and the next strategic goal deepened after the selection of elemental technologies and realization of the system. The elemental technologies were selected each time when you progressed to the next step, and the strategy became clarified and evolved during the progression. As a result, I assume that the methods of so-called “strategic deepening and selective synthesis” of elemental technologies were taken (see Figure a; modification of figure in *Synthesiology* 1(2) p.141). Is this view correct?

Answer (Kohtaro Ohba)

I think the method taken in the process of product realization could be called “strategic deepening” and “selective synthesis.” However, those were dependent on personal experiences and human contacts, and I feel it is extremely difficult to spell things out logically in a clean-cut manner to share it with others. I nevertheless have added that the feasibility study

phase was greatly affected by the “strategic deepening and selective synthesis”, including selections and rejections of some components, development of algorithms dedicated to those components, and meeting with partners to realize them, as well as the “lucky coincidences”. That is because as we blindly wandered in the dark and finally arrived at the goal, only in retrospect can we say certain places were the forks in the road. While I was groping desperately, honestly speaking, I cannot say I did any strategic decision-making even in retrospect. These points are discussed in “Chapter 2: History of twists and turns.”

2 “Valley of death”

Comment (Hideyuki Nakashima, Future University Hakodate)

There are the expression “valley of death” in several places. I think valley of death refers to the situation in which although there is technological prospect, other conditions (particularly cost) cannot be fulfilled toward product realization. It is a gap in research and development, where people do not want to get involved because it is not interesting as basic research since the principle is known, or because the development cost is too large for a company.

Answer (Kohtaro Ohba)

I think there is some difference in perception of the term “valley of death” as I understand it. The reviewer writes, “I think valley of death refers to the situation in which although there is some technological prospect, other conditions (particularly cost) cannot be fulfilled toward product realization,” but I don’t think it is the valley of death if one sees some technological prospect. Here, I use valley of death as a “place where one wanders in, loses the sense of direction of the goal much like the Forest of Aokigahara, and must make numerous twists and turns before arriving at the exit.”

Comment (Hideyuki Nakashima)

I think the part where the prospect of research cannot be seen (scene where a breakthrough is necessary) can be called the “wall,” “obstacle,” or “bottleneck.” I think you should refer to the “valley of death” in Wikipedia (Japanese version).

3 Synthetic explanation

Comment (Hideyuki Nakashima)

When we write an ordinary research paper, it is a presentation of a pathway taken in afterthought as we look back from the point the research is completed. We write as if the way to the conclusion was a straight path where we made all the choices without any doubt. The choices that were not selected or the accumulation of failures are not described in the paper. In that sense, the first manuscript was written like an ordinary paper.

Ordinary logical and analytical papers can have such structures. That is because the phenomena that must be understood are already in existence, and the purpose of a thesis is to present the pathway to the understanding. However, *Synthesiology* is for the discipline of synthesis. There may be more than one answer. The selection of certain paths amongst several possibilities itself is an important factor in synthesiology, and the descriptions of the selections are necessary. I think the content explained in the beginning of “Chapter 2: History of twists and turns” is important.

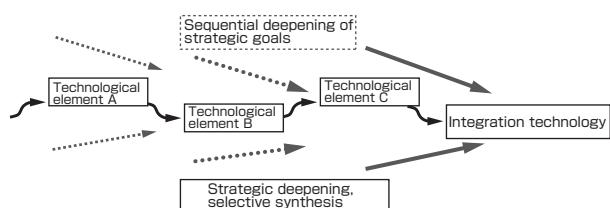


Fig. a Structure of synthesis.