Multi-beam confocal microscopy based on a custom image sensor with focal-plane pinhole array effect

Keiichiro Kagawa,¹ Min-Woong Seo,¹ Keita Yasutomi,¹ Susumu Terakawa,² and Shoji Kawahito¹

¹Research Institute of Electronics, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu, Shizuoka 432-8011, Japan
²Photon Medical Research Center, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu, Shizuoka 431-3192, Japan

Abstract: Multi-beam confocal microscopy without any physical pinhole was demonstrated. As a key device, a custom CMOS image sensor realizing a focal-plane pinhole array effect by special pixel addressing and discarding of the unwanted photocarriers was developed. The axial resolution in the confocal mode measured by FWHM for a planar mirror was 8.9 μm, which showed that the confocality has been achieved with the proposed CMOS image sensor.

©2013 Optical Society of America

OCIS codes: (040.1240) Arrays; (180.1790) Microscopy.

References and links

1. Introduction

Multi-beam confocal microscopy [1] is an effective tool for the fluorescence observation of biological tissues because of lower optical damage and faster image acquisition than those of the single-beam method [2,3]. Well-established multi-beam confocal microscopes are based on a combination of spinning microlens and pinhole array disks [4]. A fixed beam splitter or a dichroic mirror is placed between the lens array disk and the pinhole array disk for observation. As the disks rotate, the specimen is scanned in two dimensions by the multiple beams. Hence, the confocal image can be observed with an ordinary image sensor.

In this paper, we propose a new scheme to embody the multi-beam confocal microscopy without any physical pinhole and show that axial and transverse resolutions are improved in the experiment. Because spinning disks are not required, it will be possible to make the system more compact and more cost-effective. The advantage of the low optical damage is also effective in the proposed multi-beam confocal microscopy. In the proposed scheme, a dedicated complementary metal-oxide-semiconductor (CMOS) image sensor called a multi-beam confocal CMOS image sensor, which works as an electronically reconfigurable pinhole array, is a key device. CMOS image sensors [5] can integrate various kinds of analog and digital circuits, which contribute to flexible addressing [6] and processing for image compression, wide dynamic range imaging, and noise reduction [7,8] on the chip. Recently, applications of dedicated CMOS image sensors to microscopy have been explored [9–11]. With special addressing for pixel readout and control of photocarrier accumulation, what we call a focal-plane pinhole array effect is realized. In addition, by modifying only the control signals, we can switch the normal (non-confocal) and confocal imaging modes. This paper describes an optical setup for the proposed multi-beam confocal microscopy. The principle, architecture, and operation of the multi-beam confocal CMOS image sensor are also shown. Confocality of the proposed scheme is experimentally confirmed with a prototype 256 × 256-pixel CMOS image sensor designed for the illumination by 32 × 32 light beams.

![Fig. 1. Optical setup of the proposed multi-beam confocal microscope.](image)

2. Multi-beam confocal microscopy with a custom CMOS image sensor

An optical setup of the proposed system is shown in Fig. 1. Two-dimensional zigzag scan is used in our scheme to ensure the compatibility of the structure of most CMOS image sensors where pixels with a discrete structure are aligned on a two-dimensional regular grid. To achieve the confocality, the positions of every light spot and the corresponding pixel have to be accurately identical. Therefore, any sub-pixel shift, which is observed in the conventional
continuous rotational scan with the spinning disk [4], is not allowed. There are several methods to implement the zigzag multi-beam scanning, for example, by galvano mirrors or a spatial light modulator. We have selected a method of moving the microlens array mounted on a two-dimensional translation stage. The microlens array is illuminated by a collimated laser beam, which generates an array of the focused light spots. It is necessary that the light is normally incident to the array for obtaining the smallest circular spots because off-axis aberrations of microlenses are not corrected sufficiently. The advantage of this configuration is shorter optical path compared with that for two-dimensional beam steering mirrors. The spot array generated by the microlenses with the pitch of \( p \) is reduced on the observed plane, where the spot pitch becomes \( (f_{\text{obj}}/f_{\text{RD}}) p \). On the observed plane, every spot has to be diffraction-limited. The reflected, scattered, or fluorescent light from the specimen is imaged on the image sensor with the magnification ratio of \( f_{\text{mic}}/f_{\text{obj}} \). When the fill factor of the image sensor is almost 100\%, the pixel pitch should be around \( \Delta = (f_{\text{mic}}/f_{\text{obj}}) \delta \) where \( \delta \) is the spot size defined by the peak to the first minimum distance on the observed plane to obtain the confocality. An aperture stop is inserted after the collimator lens with \( f_{\text{RD}} \) to block the stray light that does not enter the pupil of the collimator lens. The peak-to-peak dislocation of the microlens array is a little less than \( p \), which is very small, for example 28 \( \mu \text{m} \) as mentioned in Section 4. Therefore, it is potentially possible to move the array at a high frequency, for example, more than hundreds of Hz in a resonant operation mode.

3. Principle and architecture of the custom CMOS image sensor

Figure 2 shows the procedure to perform the focal-plane pinhole effect on the image sensor. There are \( M \times M \) \((M = 2)\) light spots, and the image sensor is composed of \( M \times M \) blocks each of which has \( N \times N \) \((N = 4)\) pixels. The spot array is two-dimensionally scanned in a zigzag manner as shown in Fig. 2(a). When the light spots are about to enter the pixels in particular rows, all the pixels in the rows are reset to discard all photogenerated carriers at the photodiodes (Fig. 2(b)), and then accumulation begins (Fig. 2(c)). When the light spots are moving to the pixels in the next rows, the accumulation ends. Because only pixels in conjugate with the light spots are read (Fig. 2(d)) and the other photo signals are discarded by the reset operation, the pixels work as a pinhole array. As depicted in Figs. 2(e)-2(g), the procedure for the next pixels repeats.

In general, transverse and axial resolution enhancement called a confocal effect is achieved by suppressing spread defocused light with a pinhole located in the plane conjugated with the observed plane. Based on the above mechanism, the multi-beam confocal image sensor works as a pinhole array as well as a photodetector array. On the other hand, ordinary CCD and CMOS image sensors capture both of in-focus and out-of-focus light. Therefore, the confocal effect cannot be obtained without an additional appropriate pinhole array.
Fig. 2. Principle to perform a focal-plane pinhole array effect with multi-beam scanning. (a) scanning path of the light beams, (b)(e) reset, (c)(f) accumulation, and (d)(g) read.

Fig. 3. Image sensor architecture.

Figure 3 shows the block diagram of the multi-beam confocal CMOS image sensor to implement the procedure shown in Fig. 2. The image sensor consists of the pixel array which is composed of $M \times M$ blocks each of which has $N \times N$ pixels, 2-stage vertical and horizontal scanners, and correlated double sampling (CDS) circuits [7], a timing generator, and an output buffer. The pixel circuit is an ordinary 4-transistor type [12]. The position of a block and the local position of a pixel in the block are denoted by $m$ and $j$, respectively. The vertical scanner is composed of a vertical block scanner and $M$ in-block scanners, which select one (block- or pixel-) row from the top to the bottom sequentially. The vertical block scanner selects one block-row $m$, and each of the in-block scanners selects one pixel-row $j$ to turn on SEL, $m$, RST $m$, and TX $m, j$. Because these control signals are shared in a row, they are identified by vertical positions $m$ and $j$ for the block and the pixel, respectively.
As shown in Fig. 4(a), the operation begins with a discard period \( T_D \) by turning on \( \text{RST}<m=0,...,M−1,j> \) and \( \text{TX}<m=0,...,M−1,j> \) to clear the existing photocarriers. Because the image sensor works in a rolling shutter mode, every row \( <m,j> \) is sequentially reset by scanning \( m \) from 0 to \( M−1 \). The accumulation time is defined by the period between the reset and the readout (\( T_A \)). In the readout period, the photo-generated carriers are transferred to the floating diffusion node (FD) by turning on \( \text{TX}<m,j> \). Then, the output of the pixels \( <i,j> \) of blocks \( <k=0,...,M−1,m> \) are read out through the vertical signal lines \( V_{\text{SIG}}<k=0,...,M−1,i> \). There are \( N \) vertical signal lines for one block-column. The multiplexer (MUX) selects one of the \( N \) vertical signals using a set of control signals, \( \text{DEC}[0:L−1] \). The pixel outputs are denoised by the CDS circuits. Then, one of the blocks to read out is specified by the horizontal block scanner. This readout operation repeats for \( m=0 \) to \( M−1 \). The timing generator controls the horizontal block scanner, which reads \( N \) pixel
signals from the left block to the right using a clock signal, HCK. The read period is as long as $T_R$. Between the discard and read periods, a shared accumulation time ($\tau_A$) is inserted to increase the accumulation time. Figure 4(b) summarizes the output format of the image sensor. Because the output signals are shuffled, the pixels values are rearranged based on the relationship $<x, y> = <kN + i, mN + j>$ on a computer. Finally, a confocal image is reproduced. If the discard period is omitted and every pixel is reset once in a single frame, the image sensor provides a normal image without confocality.

The rolling shutter operation of the CMOS image sensor introduces a difference in the light-spot position and the accumulation time that depends on the block position. Figure 5 illustrates the light spot position based on the pixel for different block positions. A general case is shown in Fig. 5(a). The position of the light spot where the accumulation starts becomes the lower and the accumulation time, $T_A$, becomes the longer for the lower block. For a pixel in the top block $<m=0>$, accumulation begins when the light spot is situated at the top of the pixel and ends in the middle of the pixel. The accumulation time is denoted by $T_{A,\text{min}} = T_D + \tau_A$. For the middle block $<m\sim M/2>$, accumulation begins when the light spot is a little below the top of the pixel, and ends before the bottom of the pixel. For the bottom block $<m=M-1>$, it starts in the middle and ends at the bottom. The time delay based on the top block and the period of the accumulation are $\{(M-1)/M\}T_D$ and $\tau_A + \{(M-1)/M\}T_R$, respectively. Because most part of the light spot is accommodated in the pixel, obvious degradation in captured images might not occur. However, the difference of the accumulation time is undesirable. Figure 5(b) shows a specific case for $T_D = T_R$ and $\tau_A = 0$, in which there is no difference in the accumulation time. This configuration is used in Sec. 4.

The frame rate is considered as follows. The period for one light-spot position, $T_p$, is defined by

$$T_p = T_D + \tau_A + T_R. \quad (1)$$
Then, the period for one frame, \( T_f \), is given by

\[
T_f = N \cdot T_p^2.
\]  

When a unit discard time and a read time for one row are denoted by \( \tau_D \) and \( \tau_R \), respectively, \( T_D \) and \( T_R \) are equal to \( M \tau_D \) and \( M \tau_R \), respectively. \( \tau_R \) is further decomposed to the CDS operation time, \( \tau_{CD} \), and \( M \) times one-pixel read time, \( \tau_{PIX} \).

\[
\tau_R = \tau_{CD} + M \tau_{PIX}.
\]  

The frame rate of the CMOS image sensor, \( f_r \), given by \( 1/T_f \) is represented as follows:

\[
f_r = \frac{1}{N \cdot (T_D + \tau_A + T_R)}
\]

\[
= \frac{1}{N^2 \cdot (M \cdot (\tau_D + \tau_{CD} + M \cdot \tau_{PIX} + \tau_A) + \tau_A)}.
\]  

If any sufficiently fast beam steering device is available, \( f_r \) is equal to the maximum frame rate of the proposed multi-beam confocal microscope system.

To evaluate the frame rate, the actual parameters of the prototype image sensor shown in Sec. 4 are considered. When we assume \( N = 8 \), \( M = 32 \), \( \tau_D = 1.15 \) μs, \( \tau_A = 0 \) μs, \( \tau_{CD} = 4.2 \) μs, and, \( \tau_{PIX} = 0.2 \) μs, the one-position time, \( T_p \), and the one-frame time, \( T_f \), become 376 μs and 24.1 ms, respectively. Thus, the frame rate is 41.6 frames per second (fps). However, in this case, the accumulation time varies depending on the block position. \( T_{A,min} \) and \( T_{A,max} \) are 36.8 μs and 328.6 μs, respectively. When \( T_D \) is extended to be equal to \( T_R \), namely, 339.2 μs, \( T_A \) becomes constant and equals \( T_R = 339.2 \) μs. In this condition, the frame rate is 23fps.

4. Experimental results

A prototype chip of the multi-beam confocal CMOS image sensor was fabricated. The specifications and the measured characteristics are summarized in Table 1 and 2. The microphotograph of the sensor is shown in Fig. 6. \( M \) and \( N \) were 32 and 8, respectively. In the experiment, a 532 nm laser coupled with a single-mode optical fiber was used as a light source. The pitch of the microlens array was 28μm, and the shape of the elemental lens was rectangular. \( f_{BD} \), \( f_{BL} \), and \( f_{IMG} \) were 65.8 mm (Zeiss, EC Plan Neofluar 2.5 ×, NA/0.075), 20 mm (Edmund Optics, EO Plan Apo 10 ×, NA/0.28), and 141 mm, respectively. The calculated resolution of the objective lens for 532 nm was 1.16 μm. The equivalent resolution on the image sensor was about 8.2μm, which was comparable with the pixel pitch of 7.5μm.

The measured spot size defined by the peak to the first minimum distance on the observed plane was about 1.1μm, which was almost the same as that of the diffraction-limited spot. For moving the spot array, the microlens array was mounted on two-dimensional piezo stages (PI, Model P-612.2SL, closed-loop control with strain gauge sensors, maximum translation of 100μm). The piezo stages were operated with triangular and step waveforms (for fast and slow axes, respectively) to perform the zigzag scan. The monitored waveforms of the position sensors of the piezo stages are shown in Fig. 7. The proposed image sensor was demonstrated at a low frame rate of 0.17 Hz because the scanning rate depended on the maximum velocity of the two-dimensional piezo stages limited in terms of the resonant frequency and heat.
generation (horizontal and vertical scanning rates of 0.17 Hz and 1.36 Hz, respectively). When the stages return to the initial vertical or horizontal position, the closed-loop control was lost for a short time. Therefore, a blank time equal to 50% of the vertical scan of \( N \) pixels was inserted at each end of the horizontal and vertical scan. The accumulation time was 30.31 ms. The laser was operated in the CW mode in the experiment. 1951 USAF test target (Edumond optics, positive reflective type) was observed as a specimen.

### Table 1. Specifications of the prototype chip

<table>
<thead>
<tr>
<th>Technology</th>
<th>0.18μm 1-poly 4-metal CMOS image sensor process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chip size</td>
<td>5.0mm sq.</td>
</tr>
<tr>
<td>Power supply</td>
<td>1.8V (digital), 3.3V (analog)</td>
</tr>
<tr>
<td>Pixel count</td>
<td>256 × 256</td>
</tr>
<tr>
<td>Pixel architecture</td>
<td>4-transistor with a pinned photodiode</td>
</tr>
<tr>
<td>Pixel size</td>
<td>7.5μm sq.</td>
</tr>
<tr>
<td>Fill factor</td>
<td>45% (without microlens)</td>
</tr>
<tr>
<td>Sensitive area</td>
<td>1.92mm sq.</td>
</tr>
<tr>
<td>ADC resolution</td>
<td>12-bit (external)</td>
</tr>
<tr>
<td>Maximum frame rate</td>
<td>30 frames per second (normal mode),</td>
</tr>
</tbody>
</table>

### Table 2. Measured characteristics

<table>
<thead>
<tr>
<th>Photosensitivity</th>
<th>25.9ke-/lx s (with 3740K light source and IR cut filter)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel conversion gain</td>
<td>70μV/e-</td>
</tr>
<tr>
<td>RMS noise</td>
<td>11.8e- (gain of × 1), 4.9e- (× 32)</td>
</tr>
<tr>
<td>Power consumption</td>
<td>124.5mW</td>
</tr>
</tbody>
</table>

Figure 6. Photomicrograph of the prototype sensor.

Figure 7. Monitored waveforms of the position sensors of the piezo actuators. Upper: fast axis, bottom: slow axis.

Figure 8 shows an example snapshot by the prototype image sensor in the normal mode without beam scanning after the alignment. 32 × 32 light spots whose pitch was exactly 60μm or the size of one block illuminate the test chart. Note that a part of the light spots is not seen because the patterned area on the test chart is transparent.
The optical system required fine alignment so that the light spots should precisely shine only the pixels to be read out at every scanning position. The quality of the reproduced image was very sensitive especially to the rotational error around the optical axis between the image sensor and the microlens array and the magnification ratio of the observation optics. The magnification was adjusted by the imaging lens \( f_{\text{mag}} \) that was implemented as a zoom lens (Edmund Optics, 10 × Manual Telephoto Video Lens, focal length of 16-160mm). The alignment was confirmed in the normal mode in the static condition, namely without beam scanning. To visually confirm the alignment, the pixels were shuffled based on the relationship \( <x, y> = <k + iM, m + jM> \). Every segment in these images was composed of \( M \times M \) pixels that belonged to the same local position \( <i, j> \). If the alignment was perfect, only one segment became bright and uniform as shown in Fig. 9(a). Figures 9(b)-9(d) show examples with some misalignment such as the rotation of the microlens array and the magnification of the observation optics. Each misalignment gave a specific pattern on the shuffled image. When there was a rotational misalignment, the shading pattern of the segment had point symmetry around the brightest segment. For a smaller or larger magnification, the shading toward or away from the center of the bright segments appeared, respectively.

![Fig. 8. A snapshot in the normal mode.](image1)

![Fig. 9. Rearranged images with (a) perfect alignment, (b) slight rotation of the microlens array, (c) smaller magnification, and (d) larger magnification.](image2)
Confocality can be verified by improvements in transverse and axial resolutions [13]. The point-spread function (PSF) for the confocal microscopy is given by the square of the PSF for the normal (non-confocal) microscopy [14]. Therefore, in the confocal mode, the transverse PSF by full-width half-maximum (FWHM) is smaller than that of the normal mode, while its radius defined by the peak to the first minimum distance does not change, for example, as shown in Ref [13]. An improvement of the transverse resolution is verified based on an increase of the contrast of fine line-and-spaces. An improvement for the axial resolution by the confocal microscopy is very characteristic. The brightness for the out-focus area decreases rapidly, which realizes axial slicing of objects. The dependency of the brightness for the axial displacement is formulated as shown in Ref [3].

Fig. 10. Comparison of captured images along the axis. Axial displacements: (a)(d) 0 μm, (b)(e) 5 μm, (c)(f) 10 μm. (a)-(c) normal mode. (d)-(f) confocal mode.

Fig. 11. Enlarged results of Fig. 10. (a)(d) 0 μm, (b)(e) 5 μm, (c)(f) 10 μm. (a)-(c) normal mode. (d)-(f) confocal mode.
Fig. 12. Cross sections of Fig. 11: (a) A-A’ and (b) B-B’.

Figure 10 shows the captured images in the normal and confocal operation modes for three different focal displacements. In the normal mode, a LED light (PiPhotonics, HL01, peak wavelength of 525nm) was placed in front of the microlens array for incoherent illumination with almost the same NA as in the confocal mode. Unnecessary photocarriers caused by the defocus were also read out without discarding them. As expected, the brightness was almost constant in the normal mode through the focusing positions. In contrast, with the focal-plane stophole array effect in the confocal mode, the image became darker as the displacement became larger. Figure 11 enlarges the groups 8 and 9 in Fig. 10(a) and (d). Note that the brightness was modified to let all images have almost the same brightness to compare the resolution and contrast. Figure 12 shows the cross-sections of line-and-spaces for A-A’ and B-B’ indicated in Fig. 11. The brightness is not modified because these parts have almost the same brightness. The width of the line was 0.975μm, which was a little smaller than the diffraction-limited resolution (1.1μm). As shown in Figs. 11 and 12, the contrast of the line-and-spaces for the confocal mode is higher than that for the normal mode. This can be caused by shrinkage of the PSF due to the confocality.

Fig. 13. Comparison of the relationships between the axial displacement and the intensity.

Figure 13 compares the relationships between the axial displacement from the best-focus position and the normalized pixel value (or the detected light intensity) for several conditions. The objective lens for $f_{\text{obj}}$ was mounted on a one-dimensional piezo stage (Nano Control Co., Ltd., NS7210-C, maximal translation of 200 μm, resolution of 10 nm) to introduce an axial displacement. Confocal and LED show the measured characteristics for the block <12,13> in the confocal mode with the laser beam scanning and in the normal mode with the LED
illumination, respectively. Note that the position of the top left block is <0,0>. These results were given by averaged values in the same block. Because there is no pattern in the block <12,13>, this part is regarded as a planar mirror object. Therefore, the confocality can be confirmed by comparing the results in Confocal with Ideal given by the dashed curve. Ideal shows the theoretical characteristics for a planar mirror object and the ideal pinhole. Point shows the pixel value of the specific single pixel in the same block as that for confocal and LED where the center of the light spot was situated. In the measurement, the beam scanning was stopped and the image sensor worked in the normal mode. Because this situation was identical to that the defocused light was filtered by a pinhole, the measured characteristics was expected to give those in the confocal mode, which should correspond to the best results given by the proposed CMOS image sensor for the confocal mode. Instead, $N \times N$ frames are required if one whole confocal image is reproduced.

The axial resolutions by FWHM for ideal, point, and confocal were 4.4μm, 7.9μm, and 8.9μm, respectively. The axial resolution for confocal was almost the same as that for point, but was about twice as that for ideal. This was because the axial resolution was mainly determined by that for point. The deviation of the axial resolution for point from ideal could be brought by the rectangular aperture of the elemental microlenses.

5. Discussions

There are several issues on the image sensor and the optics. The periodic rectangular structure and gradation in block were observed in Fig. 9, which were caused by slight misalignment in the microlens array and mismatch between the scanning of the image sensor and the light spots, respectively. When the center of light spot is displaced from the center of pixel, the detected power is reduced. Because this error is common in each block, the block-like structure is observed. To solve this problem, fabrication error of the microlens has to be reduced or the microlens array should be replaced by a diffractive optical element (DOE). Although DOEs have strong wavelength dependence, it is suitable for periodic light spots because such random displacement is not generated. If the variation in brightness is not very large, it can be alleviated by simple digital processing like shading correction. For example, in preparation, the image of a perfect mirror is captured as a reference image to show the maximum signal intensity for every pixel. Then, captured images of specimens are normalized by the reference image.

In the prototype system, light spots move continuously in vertical direction but stepwise in the horizontal direction. Due to the blurring effect, the vertical resolution can be worse than the horizontal resolution. However, this speculation is not supported by the results in Figs. 12(a) and 12(b). There is almost no obvious difference between the two axes, or the vertical resolution seems a little better than the horizontal resolution. It can be because the sampling theorem was not satisfied for these fine line-and-spaces, so that the contrast is affected by only a slight displacement of the test chart. Further analysis and experimental verification are necessary to clarify the horizontal and vertical resolutions in the proposed scheme.

Another issue is to achieve simultaneous reset of the pixels at the same local position. The architecture of the vertical scanners and the pixel should be modified to let the multiple rows reset and transfer the accumulated photocarriers at the same time, whereas they are reset and read out row by row in the prototype chip. It can be realized by the global shutter technology [15].

Increasing the frame rate is also an important issue for the proposed confocal microscope system. The frame rate can be increased with fast column analog-to-digital converter technology up to thousands of frames per second [16]. However, the beam steering device limits the frame rate. It might be effective to investigate non-mechanical beam steering methods, for example, an application of optical convolution using a two-dimensional vertical-cavity surface-emitting diode (VCSEL) array [17]. As shown in Ref [17], an arbitrary optical pattern generated by a DOE can be shifted depending on the position of the light source.
this configuration, the switching speed of the pattern shift is determined by that of the VCSEL array, which is much faster than 100MHz.

Selection of $N$ or the number of pixels in one block depends on the tradeoffs among the background signal level, the power of the excitation light, and the operation frequency of the scanning device. In single-photon fluorescence microscopy, reduction of $N$ or increase of the number of the light spots will increase the background signal level due to the crosstalk between the adjacent light spots or their defocused lights. In many applications that require high contrast, it is supposed that large $N$, for example 16 or 32, is more effective. However, it will lead to a higher scanning rate for the fast axis (vertical axis in this paper), which may cause slower frame rate. In addition, for two- or multi-photon fluorescence microscopy, the minimal light intensity to evoke the nonlinear effect determines the minimal $N$ or the maximal number of the light spots ($M \times M$). $N$ should be carefully selected for each application.

6. Conclusion

In conclusion, multi-beam confocal microscopy with a custom-designed CMOS image sensor performing a focal-plane pinhole array effect was demonstrated, and the confocality by the proposed special pixel addressing and discarding of the photocarriers were experimentally confirmed. The axial resolution in the confocal mode by FWHM is 8.9 $\mu$m (ideally 4.4 $\mu$m), which showed that the confocality was achieved with the proposed CMOS image sensor.

Acknowledgments

This work was supported by SENTAN, JST. The authors are grateful to Prof. Kawata and Prof. Inami at Shizuoka University for valuable discussions and kind supports.