

Deep Learning Machine Vision System with High Object Recognition Rate using Multiple-Exposure Image Sensing Method

Min-Jun Park¹ and Hyeon-June Kim^{1,*}

Abstract

In this study, we propose a machine vision system with a high object recognition rate. By utilizing a multiple-exposure image sensing technique, the proposed deep learning-based machine vision system can cover a wide light intensity range without further learning processes on the various light intensity range. If the proposed machine vision system fails to recognize object features, the system operates in a multiple-exposure sensing mode and detects the target object that is blocked in the near dark or bright region. Furthermore, short- and long-exposure images from the multiple-exposure sensing mode are synthesized to obtain accurate object feature information. That results in the generation of a wide dynamic range of image information. Even with the object recognition resources for the deep learning process with a light intensity range of only 23 dB, the prototype machine vision system with the multiple-exposure imaging method demonstrated an object recognition performance with a light intensity range of up to 96 dB.

Keywords : Deep learning, object recognition, machine vision system, multiple exposure, wide dynamic range.

1. INTRODUCTION

Recently, various artificial intelligence (AI) applications have been widely implemented in commercial industries according to their requirements [1-3]. AI technology is the science behind machine intelligence. Traditional algorithms for a machine system set specific rules that define the output for each input case, whereas AI algorithms build their own rule systems to achieve better decisions for all cases. Therefore, machines with AI algorithms can self-solve tasks that were once entirely dependent on humans.

The advent of deep learning has strengthened the practicality of machine learning and expanded the realm of AI. As a sub-concept of machine learning, deep learning has been developed using artificial neural networks inspired by the structure and function of the brain [4]. Deep learning algorithms present the ability to learn and improve themselves based on observations and data provided;

consequently, better decisions are achieved in each case without human intervention and assistance. Many deep learning algorithms [4,5] have been studied to enhance the efficiency of the learning process. In particular, the convolutional neural network (CNN) algorithm [6] has achieved significant progress in image recognition, which extracts and learns object features in multiple steps to generate a feature map. The performance of the CNN algorithm depends on the amount of data learned in various conditions. It is prone to making incorrect decisions and errors in unexpected conditions that are not learned.

In machine vision systems using deep learning algorithms, one or more image sensors are used to recognize object features with various edge detection methods [7-10]. Owing to the operating characteristics of the image sensor, the target object and its features are represented differently according to the light intensity condition. Thus, even with embedding advanced deep learning algorithms, the machine vision system cannot provide the right object features under conditions of unlearned light intensity. That is, the overall recognition performance is limited by the image sensors. For this reason, when we look closely to the imaging mechanism, the way to further improve object recognition performances may be obtained in a wider scope; if the image sensor can represent the target object regardless of the light condition surrounding the target object, the object recognition performance would be effectively increased even without the deep learning process for generating a feature map for all light

¹ School of Electrical Information Communication Engineering, Kangwon National University, 346 Jungang-ro, Samcheok-si, Gangwon-do 25913, Republic of Korea

*Corresponding author: Hyeonjuekkim@kangwon.ac.kr
(Received: Mar. 17, 2021, Revised: Mar. 27, 2021, Accepted: Mar. 29, 2021)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

intensities.

In this study, we propose an object recognition machine vision system based on deep learning with adopting, as a well-known technique in CISs, multiple-exposure sensing method. By effectively applying the multiple-exposure sensing technique [11-13], the prototype machine vision system increases the object recognition rate without deep learning information in high and low illumination conditions. If the proposed machine vision system fails to recognize the target object, the image sensor operates in a multiple-exposure sensing mode to detect the object that is blocked in the near dark or bright region. If the object is found, the captured images in the multiple-exposure mode are synthesized for a wide dynamic range of image information [14]. Accordingly, even without deep learning data in near dark or bright region, the object recognition rate can be effectively improved resulting in reducing the cost of data processing and computing resources.

The remained of this paper is organized as follows. Section 2 describes the proposed machine vision system applying the multiple-exposure sensing method. The experimental results and discussions of the prototype system are presented in Section 3, finally, the conclusions presented in Section 4.

2. MACHINE VISION SYSTEM

2.1 Multiple exposure sensing method

In a machine vision system, the performance of an image sensor determines the overall performance of the system because the image sensor functions as the machine’s eyes. One of the key factors in the performance of the image sensor is the dynamic range, which refers to the range in which the image sensor can capture the brightest and darkest areas of an image. That is

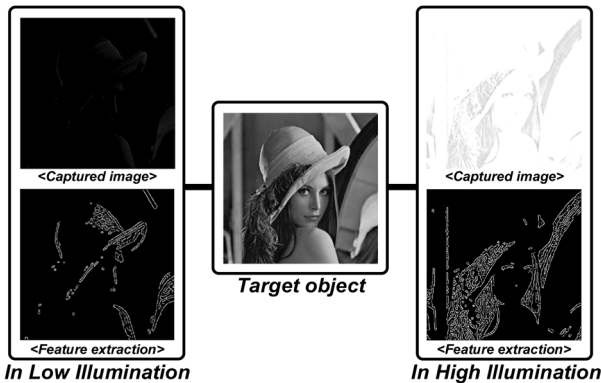


Fig. 1. Effect of light intensity.

directly related to the expressible light intensity range of the system. In particular, in a machine vision system, the edge information from the captured image is mainly used to recognize objects in various ways. However, the target object can be represented differently depending on the amount of light surrounding it, which results in different edge information. Thus, in machine vision applications, the light intensity around an object

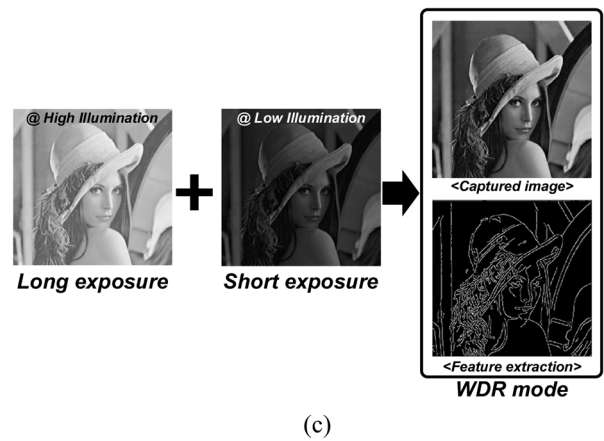
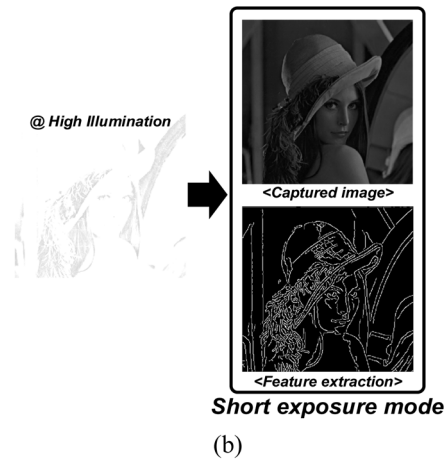
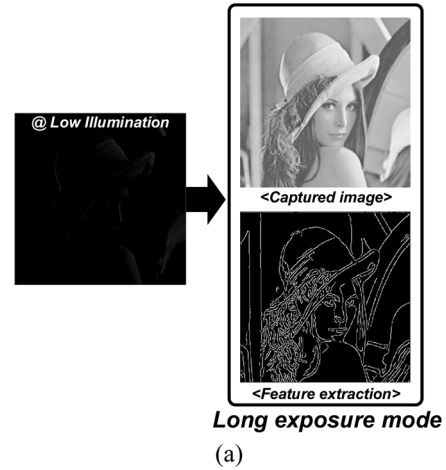


Fig. 2. Conventional multiple-exposure sensing technique.

is an important environmental factor for object recognition.

Fig. 1 is the simulation results with an 8-bit sample image in VGA pixel resolution, which shows the effect of the edge information on the light intensity around the object. For the same object, the edge information is not clearly represented under the low and high illumination conditions. That implies that the object in a scene with extremely dark or bright areas is almost invisible in machine vision. However, even in near-dark and bright conditions, as shown in Fig. 2 (a) and (b), it is possible to extract a clear image with its edge information by adjusting the exposure time of the image sensor. A multiple-exposure sensing (MES) method is commonly used to expand the dynamic range of image sensor for applications. The MES method is based on the property that the brightness of the image is proportional to the exposure time; a long exposure time is applied to prominently represent a dark image, and conversely, a short exposure time is applied to represent a bright image. Because the sharpness of the object varies according to the exposure time, the object features can be extracted by applying multiple sensing with different exposure times. In addition, as shown in Fig. 2 (c), wide dynamic range (WDR) images can be obtained by synthesizing the images captured at different exposure times, which is utilized for accurate feature extraction depending on the situation. In this work, with applying MES techniques to a prototype object recognition system based on a deep learning algorithm, we attempted to verify the performance enhancement of the object recognition.

2.2 Object feature extraction algorithm

Fig. 3 depicts the operational algorithm of the proposed machine vision system. The proposed system adopts three sensing modes to effectively obtain the object features: normal mode, MES mode, and WDR mode. The system operates as follows. First, in the normal mode, the prototype system seeks to recognize object features based on edge information. If recognition of the

target object fails, the image sensor operates in the MES mode to observe whether there are objects obscured by dark or bright light regions; subsequently, the proposed system sequentially extracts the target object features through multiple-exposure sensing images according to different predetermined exposure times. In the MES mode, the image sensor automatically changes the exposure time and can be flexibly applied depending on the environmental conditions. In the WDR mode, the images captured in the MES mode are synthesized to obtain object features more clearly as in [7]. In contrast, to the typical operating algorithm for machine vision systems, the proposed machine vision algorithm can effectively improve the object recognition performance even without deep learning data on high and low illumination conditions; thus, a reduced cost of data processing and computing resources can be realized.

2.3 Development of the prototype machine vision system

Fig. 4 shows the block diagram of the prototype machine vision system, which comprises of a camera module, Raspberry Pi 4, and 7-inch touchscreen display. For object sensing, a Sony IMX 219 PQ CMOS image sensor in a fixed-focus module was adopted in the Raspberry Pi camera module V2. The maximum frame rate was 30 frames per second, and a power supply of 1.8 V was used for optimum power consumption. In this study, as a single-core computer, Raspberry Pi 4 plays a central role in the prototype machine vision system. It is a hardware device that can handle the computational memory through a Python-based method. To display the processing results in real time, a 7-inch touchscreen display was selected by considering the size of the system exterior to be mounted. As a deep learning libraries, the Tensor flow model (SSD_mobilenet) for the deep learning process was selected because of its relatively high accuracy and processing speed. Using an open-source computer vision library (open CV),

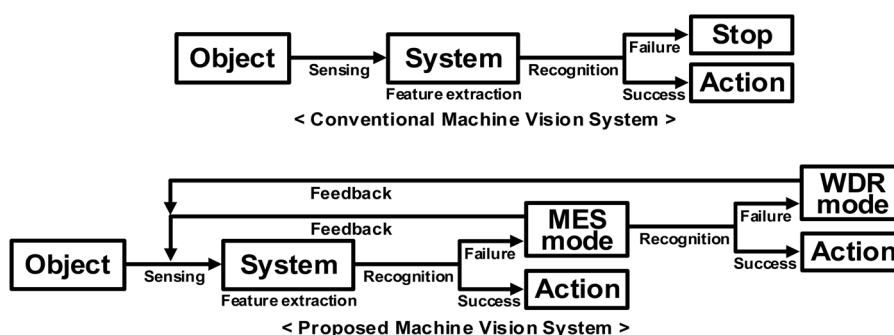


Fig. 3. Operational algorithm of the proposed machine vision system.

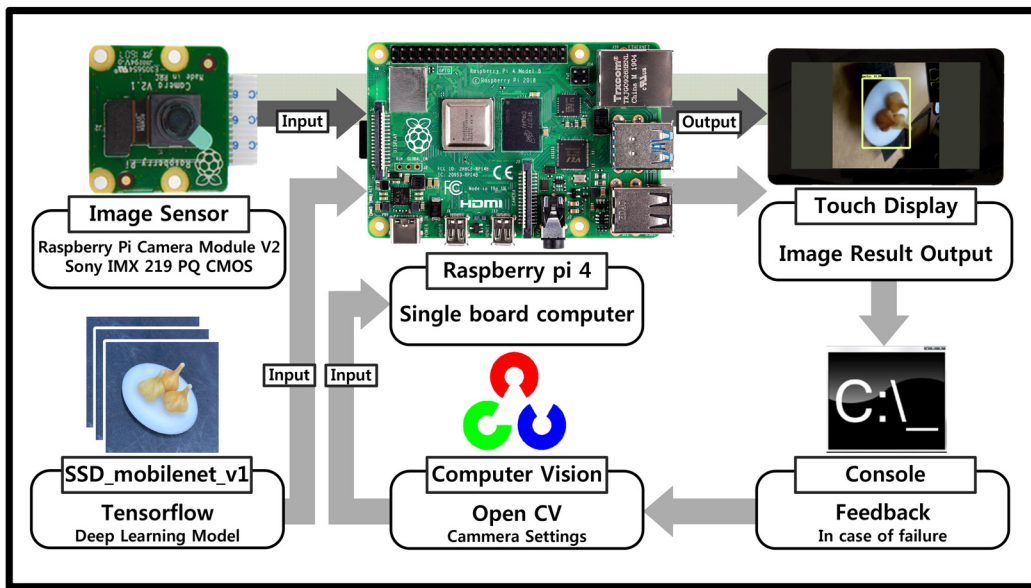


Fig. 4. Block diagram of the prototype machine vision system.

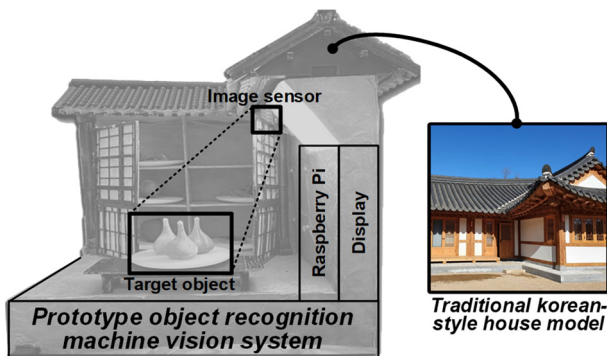


Fig. 5. Prototype object recognition machine vision model.



Model	Turbo	High	Low
ANSI lumens (LM)	3200	400	15
Beam intensity (cd)	20022	2550	100
Beam distance (m)	283	101	20

Fig. 6. Light source for generating wide light intensity range.

the prototype system can control the internal settings of the camera module and the exposure time.

3. EXPERIMENTAL RESULTS

Fig. 5 shows the prototype machine vision system model that provides information related to the target object. To verify the performance of the prototype system, deep learning was conducted using training and testing images in the ratio of 8:2. After supervised learning for image classification, the number of training repetitions was set to 200,000. To verify the object recognition performance according to the light intensity, the light intensity range of the light source should be larger than the dynamic range of the image sensor in the machine vision system. As shown in Fig. 6, a commercial light source (XT11X) was used

to generate a wide light intensity range of 15–3200 lm. Here, the maximum dynamic range of the image sensor was approximately 60 dB. The various exposure times of the image sensor can be applied by adjusting its frame rate. When the frame rate is reduced, the ratio of the illumination time for imaging pixels is proportionally increased in the total image sensing time, resulting in long-exposure image sensing; conversely, when the frame rate is increased, the ratio of the illumination time is reduced, resulting in short-exposure image sensing.

Fig. 7 shows the measurement results of the prototype machine vision system applying the multiple-exposure imaging method. Compared to the conventional system operating only in the normal mode, the prototype system operates in the MES mode, and perceives the target object with a recognition rate of over 87% and 77% in the near-dark (under 2 lux) and bright (over 130,000

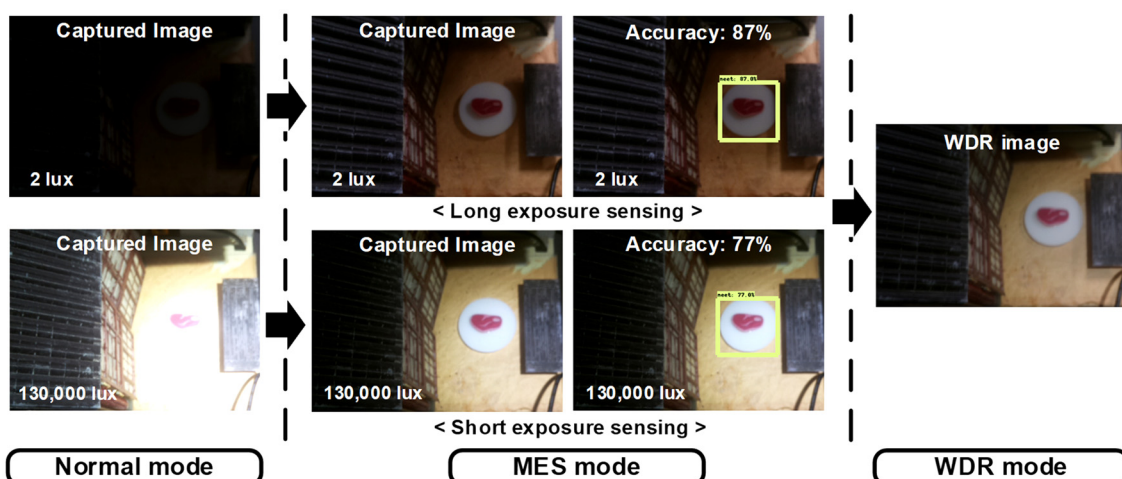


Fig. 7. Measurement results of the prototype machine vision system with multiple exposure imaging method.

lux) conditions, respectively. The amount of light intensity at a distance of approximately 1 cm from the light source was measured using an illuminometer (TES-1335). In addition, the WDR image and its edge information are generated by synthesizing short- and long-exposure images captured in the MES mode. Thus, both dark and bright areas can be expressed regardless of the performance of the image sensors in the prototype system.

Fig. 8 shows a graph of the amount of recognizable illuminance. Although deep learning was performed only with sample images in the light intensity range of 70 lux to 1,000 lux, the prototype machine vision system with the multiple-exposure image sensing method demonstrated a high object recognition performance from 2 lux up to a light intensity of over 140,000 lux. Here, the visible light intensity range of the machine vision system can be calculated as in [8]:

$$dynamic\ range(dB) = 20 \log\left(\frac{max.light\ intensity}{min.light\ intensity}\right) \quad (1)$$

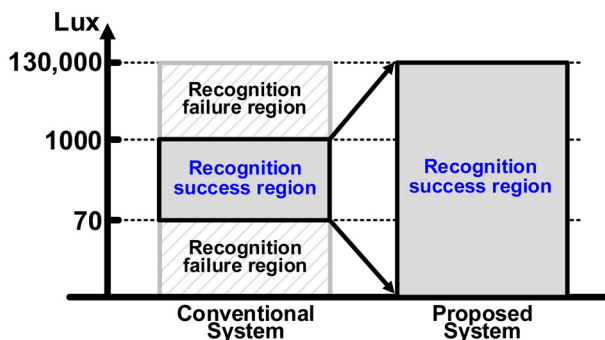


Fig. 8. Measured object recognition performance.

The measured accuracy of object recognition according to the light intensity is shown in Fig. 9. Compared to the typical machine vision system, when the MES method is applied, the better object recognition rate with high accuracy is verified in a wider light intensity range.

In general, the recognition rate and accuracy of the machine vision system depend on the amount of deep learning data and the performance of the image sensor. Nevertheless, even with the object recognition resources (based on deep learning data) with a light intensity range of only 23 dB (the light range of 70–1,000 lux), the prototype machine vision system with multiple-exposure image sensing method was verified to improve the object recognition performance with a light intensity of up to 96 dB (the light range of 2–140,000 lux). Moreover, by synthesizing the images captured in the MES mode, the prototype machine vision system provides a wide dynamic range performance beyond the dynamic range of the image sensor. Thus, this is a meaningful result because the prototype system can reduce the amount of deep learning data and represent object features beyond the dynamic range of the image sensor. The prototype machine vision system with the multiple-exposure image sensing method has

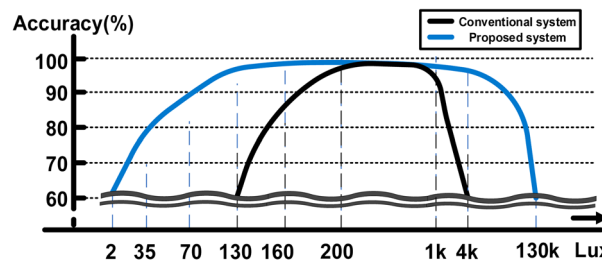


Fig. 9. Measured accuracy of object recognition according to the light intensity.

advantages in terms of the cost of data processing and computing resources.

4. CONCLUSIONS

This work presents a prototype machine vision system with a high object recognition performance. By applying the multiple-exposure sensing method, the prototype machine vision system can recognize the target object in a wide light intensity range without conducting the deep learning process under all light intensity conditions. Even in near-dark and bright conditions where the object feature is invisible, the prototype system shows a high object recognition performance. Thus, the proposed object recognition technique can be utilized to enhance the application of a machine vision system based on deep learning.

ACKNOWLEDGMENT

The study was supported by 2020 Research Grant from Kangwon National University. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R111A3074020).

REFERENCE

- [1] M. Minsky, "Steps toward Artificial Intelligence", *Proc. of IEEE*, Vol. 49, No. 1, pp. 8-30, 1961.
- [2] K. Robert, "Artificial Intelligence and Human Thinking", *International Joint Conference on Artificial Intelligence*, Vol. 1, No. 11, pp. 11-16, 2011.
- [3] S. Russell and P. Norvig, *Artificial Intelligence: a Modern Approach*, Prentice Hall Press, US, pp. 1-932, 2009.
- [4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", *Nature*, Vol. 521, pp. 436-444, 2015.
- [5] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, the Massachusetts Institute of Technology Press, Cambridge, US, 2016.
- [6] P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis", *Proc. of IEEE Conf. on August. 7th International Conference on Document Analysis and Recognition*, pp. 958-963, Edinburgh, UK, 2003.
- [7] H. Kim, "CMOS image sensor for wide dynamic range feature extraction in machine vision", *The Institution of Engineering and Technology. IET*, Vol. 57, No. 5, pp. 206-208, 2020.
- [8] H. Kim, S. Hwang, J. Chung, J. Park, and S. Ryu, "A Dual-Imaging Speed-Enhanced CMOS Image Sensor for Real-Time Edge Image Extraction", *Journal of Solid-State Circuits*, Vol. 52, No. 9, pp. 2488-2497, 2017.
- [9] H. Kim, S. Hwang, J. Kwon, D. Jin, B. Choi, S. Lee, J. Park, J. Shin, and S. Ryu, "A Delta-Readout Scheme for Low-Power CMOS Image Sensors With Multi-Column-Parallel SAR ADCs", *Journal of Solid-State Circuits*, Vol. 51, No. 10, pp. 2262-2273, 2016.
- [10] H. Kim, "A Sun-Tracking CMOS Image Sensor with Black-Sun Readout Scheme", *Trans. Electron Devices*, Vol. 68, No. 3, pp. 1115-1120, 2021.
- [11] M. Mase, S. Kawahito, M. Sasaki, Y. Wakamori, and M. Furuta, "A wide dynamic range CMOS image sensor with multiple exposure-time signal outputs and 12-bit column-parallel cyclic A/D converters", *Journal of Solid-State Circuits*, Vol. 40, No. 12, pp. 2787-2795, 2005.
- [12] J. Chun, H. Jung, and C. Kyung, "Dynamic-Range Widening in a CMOS Image Sensor Through Exposure Control Over a Dual-Photodiode Pixel", *Trans. Electron Devices*, Vol. 56, No. 12, pp. 3000-3008, 2009.
- [13] D. Stoppa, A. Simoni, A. Baschiroto, M. Vatteroni, and A. Sartori, "A 120-dB Dynamic Range CMOS Image Sensor with Programmable Power Responsivity", *Proc. of IEEE Conf. the 32nd European Solid-State Circuits.*, pp. 420-423, Montreux, ES, 2006.
- [14] J. Park, S. Aoyama, T. Watanabe, T. Akahori, T. Kosugi, K. Isobe, Y. Kaneko, Z. Liu, K. Muramatsu, T. Matsuyama, and S. Kawahito, "A 0.1e- vertical FPN 4.7e- read noise 71dB DR CMOS image sensor with 13b column-parallel single-ended cyclic ADCs", *IEEE Conf. International Solid-State Circuits*, pp. 268-269a, San Francisco, US, 2009.