In this study, we propose a multi-thread artificial intelligence (AI) camera system that can simultaneously recognize remote objects in desired multiple areas of interest (AOIs), which are distributed in a wide field of view (FOV) by using single image sensor. The proposed multi-thread AI camera consists of an ultrafast active vision system and a convolutional neural network (CNN)-based ultrafast object recognition system. The ultrafast active vision system can function as multiple virtual cameras with high spatial resolution by synchronizing exposure of a high-speed camera and movement of an ultrafast two-axis mirror device at hundreds of hertz, and the CNN-based ultrafast object recognition system simultaneously recognizes the acquired high-frame-rate images in real time. The desired AOIs for monitoring can be automatically determined after rapidly scanning pre-placed visual anchors in the wide FOV at hundreds of fps with object recognition. The effectiveness of the proposed multi-thread AI camera system was demonstrated by conducting several wide area monitoring experiments on quick response (QR) codes and persons in nature spacious scene such as meeting room, which was formerly too wide for a single still camera with wide angle lens to simultaneously acquire clear images.

**Keywords:** visual monitoring, convolutional neural network, object recognition, high-speed vision, multithread viewpoint control

1. Introduction

Intelligent visual monitoring is an active research topic in computer vision area, and it is used to automatically capture, recognize, and track objects or events of interest in real-world situations, such as traffic monitoring [1], public security [2], and attendance check [3]. Besides, it is becoming increasingly important for robotics applications, such as visual odometry [4, 5]. Recently, with the increasing demand for reliable visual monitoring in wide area scenarios, such as stadiums, squares, and theaters, numerous wide area surveillance systems have been developed to increase the image quality for better object recognition in wider FOV. Since spatial resolution of single image sensor is limited, most of conventional wide area surveillance systems increase number of image sensor or change orientation of camera axis to expand visual surveillance coverage without degrading image resolution.

Numerous high-speed vision systems [6, 7] which can acquire images at hundreds of frames per second (fps) or higher have been proposed. Many high-speed active vision systems which consist of a high-speed camera operating at hundreds of fps or higher, and a servo motor system have been reported to expand visible area without losing image resolution. Many image processing algorithms, such as pixel-level short-time Fourier transform (STFT) [8], optical flow [9], Camshift [10], fast trackers based on face detection [11], and CNN-based object recognition [12] have been accelerated by field-programmable gate array (FPGA) board or graphic processing units (GPUs), and implemented on above-mentioned systems for variable applications, such as multi-copter tracking [13], omnidirectional visual monitoring [14], microscopic-cell imaging [15], and vibration analysis [16]. Recently, an ultrafast gaze control system which consists of a high-speed camera and a shutter-synchronized Galvano mirrors has been developed for simulating multiple virtual pan-tilt cameras by switching the camera viewpoint at hundreds of times in a second; and has been applied to many applications such as large structure health monitoring [17], monocular stereo sensing [18]. Compared to conventional PTZ camera network, the ultrafast gaze control system can flexibly designate position and number of monitoring scope in wide area without increasing hardware cost.
In this study, we propose a multi-thread AI camera system that can simultaneously detect and recognize remote objects in desired AOIs, which are distributed in a wide FOV by using an ultrafast gaze control system. The remainder of this article is organized as follows. In Section 2, we introduce conventional hardware and software-based approaches for acquiring high-resolution wide-FOV images. In Section 3, we propose the concept of the multi-thread AI camera system and describe its hardware specifications. In Section 4, we describe the viewpoint control policy and multi-GPU-based acceleration of object recognition in the proposed system. To verify its performance, in Section 5, we present remote object monitoring result for persons in wide-area environment.

2. Related Works

2.1. Hardware-Based FOV Extension

Optical components such as wide-angle lenses [19,20], convex mirrors with shapes of hyperboloid [21], ellipsoid [22], paraboloid [23], and conic [24], and their integrations [25, 26] have been extensively applied in camera systems for wide-area visual monitoring applications, which do not require high spatial resolution images for remote object. They offer an extended view angle with single image, but always suffer from geometrical distortion issues, which result in deformation of the object shape besides resolution degradation. Many distortion model estimation approaches, which can be categorized into geometric projection [27], image feature [28], and calibration pattern-based [29] methods, have been proposed to correct the distortions. However, it is difficult to recover high-quality images after correction due to the spatial-resolution-loss especially in peripheral region.

To enable remote observation of desired areas in wide-FOV with high-spatial resolution images, many pan-tilt-zoom (PTZ) cameras which provide high maneuverability and optical zooming ability have been developed for various applications, such as homography estimation [30, 31], broadcasting [32], mobile robot [33], and wide-area visual surveillance [34]. Meanwhile, many algorithms for object detection or tracking [35,36] have been provided to realize automatic visual servo tracking for remote object with PTZ camera. However, PTZ movements are relatively slow and view angle of PTZ camera always become narrow especially when using high focal length, so that it is difficult to simultaneously monitor multiple objects or events in wide area with single PTZ camera.

Since above-mentioned FOV expansion methods based on single camera have intrinsic bottleneck between image resolution for remote objects and monitoring coverage, multi-camera vision systems which expand FOV without losing image resolution by increasing number of image sensors have been intensively studied for wide-area visual surveillance systems. In order to reduce system cost under the condition of achieving seamless monitoring coverage, conducting global calibration for multiple cameras is important to narrow overlapping field of views between different cameras. Many non-overlapping calibration methods have been proposed to estimate extrinsic camera parameters, such as tracking- and mirror-based methods [37–39] for stationary camera cluster, and SLAM- and marker-based methods [40–42] for movable cameras. Besides, many active camera systems mounted with long-focal-length lens have been developed for mobile robots that can search and acquire high resolution images of small patterns in the distance for detecting and recognizing artificial landmarks in wide area to realize operations of path planning [43]. However, achieving seamless coverage also requires tremendous number of camera systems especially when monitoring a wide-area environment using long-focal-length lens, and intelligently excludes non-interested regions for saving cost.

Since above-mentioned hardware-based approaches for FOV extension adopt standard video format signal with dozens of frames per second such as PAL or NTFS, visual scanning takes a long time to cover a wide area seamlessly with actuators, and they also have intrinsic bottleneck between image resolution for remote objects and monitoring coverage.

2.2. Software-Based Wide-FOV Image Refinement

Besides above-mentioned approaches which increase hardware cost, numerous software-based methods have been proposed to increase the object recognizability in wide area when the image resolution is insufficient. Super-resolution approaches have been conducted on enhancing the character details of single low-resolution wide-angle image; existing approaches can be classified into two categories: model- and deep learning-based methods.

To regain pattern details more flexibly, advanced model-based super-resolution approaches exploiting more complex image priors in multiple viewpoints such as non-local similarity [44] and sparsity prior [45] have been proposed. Although these methods have shown relatively superior reconstruct performance, they still suffer from computationally heavy problems.

Recently, deep learning based advanced super-resolution models have been intensively studies for single image reconstruction, and several super-resolution-oriented convolutional networks (such as SRCNN [46], DRCNN [47]) and generative adversarial networks (such as SRGBAN [48]) have achieved the state-of-the-art performance.

However, while cutting edge software-based methods show significant performance in low-resolution image reconstruction, minimum apparent cues of pattern details are required, such that measurable distance is significantly limited.
3. Multi-Thread AI Camera System

3.1. Concept

As mentioned in Section 1, conventional wide area visual monitoring methods based on optical elements, camera network, pan-tilt actuators, or super resolution approaches cannot simultaneously acquire high spatial resolution images of remote multiple AOIs owing to the inevitable trade-off between monitoring range and image quality. Besides, a large proportion of uninterested regions are normally included in FOV images where pixels of rectangular array are always not fully utilized in wide area visual monitoring applications.

We propose a multi-thread AI camera system that can simultaneously detect and recognize remote objects in desired multiple areas with high spatial resolution images. In this system, optically zoomed images of different viewpoints can be acquired at hundreds of fps by using single ultrafast pan-tilt active vision system. The ultrafast active vision system can function as multiple virtual pan-tilt cameras with long-focal-length-lenses by synchronizing a high-speed camera and a 2 degrees-of-freedom (DOF) ultrafast Galvano-mirror system that can switch hundreds of views per second. As depicted in Fig. 1, gaze of the ultrafast pan-tilt active vision system rapidly switches between the AOIs which include pre-allocated visual landmarks, so as to simultaneously acquire high resolution videos around each visual landmarks. Multithread gaze control for multiple virtual cameras is illustrated in Fig. 2, where pan-tilt mirror actuation and image acquisition processes are included in each frame interval of high-speed camera. Here, to avoid the motion blur which occurs during the mirrors’ movement, high-speed camera exposes only when the mirrors are completely stopped between adjacent pan-tilt mirror actuation processes. The simultaneously acquired videos are detected and recognized by a high-performance deep learning server which is accelerated by implementing multi-GPU-based parallel processing.

Clearly different from existing intelligent wide area visual surveillance systems which are based on PTZ camera or multicamera network, the proposed multi-thread AI camera system can simultaneously detect and recognize remote object with high resolution images of different AOIs in wide area by using single imaging system. The viewpoint positions can be flexibly modified by moving the visual landmarks without any electro-mechanical delay, and it also enables simple calibration by adjusting the viewpoints of multiple virtual cameras precisely.

3.2. Hardware Configuration

We developed a multi-thread AI camera system that can simultaneously detect and recognize remote objects from high resolution videos captured by multiple virtual cameras. The system consists of an ultrafast pan-tilt camera and its control system, and a high-performance deep learning server equipped with multiple GPUs. We depict the (a) configuration, (b) an overview, and (c) geometry of the proposed multi-thread AI camera system in Fig. 3. The ultrafast pan-tilt camera consists of a two-axis pan-tilt Galvano-mirror system (6240H, Cambridge Technology, US), a high-speed CMOS camera head (DMK37BUX287, Image Source, Germany) and a personal computer (PC) for controlling the pan-tilt Galvano-mirror system and high frame rate image acquisition. The Galvano-mirrors consists of a pan mirror (27 × 18 mm²) and a tilt mirror (36 × 27 mm²), whose maximum scan angles are ±20° and ±10°, respectively, and their rotational axes are mutually perpendicular with the shortest distance of 18 mm. Angles of the two mirrors are controlled with input voltage signals which are converted from digital commands via an AD/DA board (PEX-361116, Interface Corp., Japan). The high-speed CMOS camera, whose sensor and pixel sizes were 4.96 × 3.72 mm and 6.9 × 6.9 μm can capture and transfer 740 × 520 8-bit grayscale images at 500 fps to PC via USB 3.1 interface. Axis of the camera is parallel to tilt-mirror-axis and crosses the pan-mirror-axis, and it is set as 24 mm in front of the center of the
Jiang, M. et al.

4. Simultaneous Object Recognition Using Multi-Thread Active Vision System

We introduce the simultaneous object recognition for multiple AOIs by using multi-thread active vision system. It involves two processes: (a) AOIs determination by searching pre-allocated visual landmarks, and (b) multiple virtual cameras with object recognition realized by rapidly switching viewpoint within the selected AOIs at 500 Hz. Acquired images of each virtual camera are simultaneously recognized by CNNs in real time by implementing multi-GPU-based acceleration.

4.1. AOIs Determination Using Visual Landmarks

4.1.1. Searching Path

We describe the process of AOIs determination in wide area by searching pre-allocated landmarks. The pan and tilt angles of Galvano-mirror system are rapidly controlled to acquire high-resolution local images which cover wide-angle panoramic view. To maximally avoid long-range jumping in mirrors control, scanning path of viewpoint is designed as depicted in Fig. 4(a), and the corresponding stride angles of the pan and tilt mirror are set as half of horizontal and vertical FOV of camera lens to seamlessly cover the wide-angle panoramic area. The time intervals of input HFR images are denoted as $\delta t$.

The column and row number of scanning viewpoint, $i$ and $j$ at time $t = k\delta t$ ($k = 0, 1, 2, 3, \ldots$) are calculated as

$$i(t) = \left\lfloor \frac{k}{M} \right\rfloor$$

$$j(t) = \begin{cases} k - \left\lfloor \frac{k}{M} \right\rfloor \cdot M & \text{if } \left\lfloor \frac{k}{M} \right\rfloor \text{ is even}, \\ \left( k - \left\lfloor \frac{k}{M} \right\rfloor \right) \cdot M - k - 1 & \text{if } \left\lfloor \frac{k}{M} \right\rfloor + 1 \text{ is even} \end{cases}$$

Here $\lfloor \cdot \rfloor$ denotes floor function, $M$ and $N$ denote column and row number of the scanning viewpoint in wide-angle panoramic view, and it takes $M \cdot N \cdot \delta t$ [s] to complete a searching cycle. The viewpoint $(0, 0)$ is the left-top corner of the wide-angle panoramic view and searching, and the corresponding stride angles of the pan and tilt mirror are set as half of horizontal and vertical FOV of camera lens to seamlessly cover the wide-angle panoramic area.

4.1.2. Visual Landmark Recognition

The visual landmark candidates estimated from an input image for searching $I_M(t)$ at time $t = k \cdot \delta t$ ($k = 0, 1, 2, 3, \ldots$) are expressed as follows:

$$\mathcal{D}_L(I_M(t)) = \{d_{1}^{L}(t), d_{2}^{L}(t), \ldots, d_{L}^{L}(t)\}$$

where $\mathcal{D}_L$ denotes an operator of CNN-based visual landmark detection. For the $i$-th detected visual landmark
Multi-Thread AI Cameras Using High-Speed Active Vision System

\[ d_k^i(t) = \{ x^i(t), y^i(t), w^i(t), h^i(t), \theta^i(t) \}, \ldots (4) \]

where \( x^i(t), y^i(t) \) denote x and y centroid coordinate, \( w^i(t) \) and \( h^i(t) \) denote width and height, and \( \theta^i(t) \) denotes detection confidence of \( i \)-th detected visual landmark at time \( t \). The top-scored \( i \)-th candidate with detection confidence would be selected as the visual landmark of the current scanning viewpoint by inspecting detection confidences as follows:

\[ i = \arg \max_{S^i(t) \geq \theta} p^i(t). \]........... (5)

Here, \( S^i(t) \) denotes size of the \( i \)-th detected bounding box at time \( t \) as

\[ S^i(t) = w^i(t) \times h^i(t), \ldots (6) \]

and the threshold \( \theta \) is set to ignore mis-detected small arbitrary patterns in the environment for better system robustness. We adopt specially designed quick response (QR) code as visual landmark in the proposed system.

**4.1.3. Visual Landmark Alignment**

As depicted in Fig. 4(b), positions of the recognized visual landmarks are always deviated from desired position in viewpoints in searching stage, so that local compensation both in pan and tilt directions is required for each visual-landmark-detected viewpoint to monitor desired AOIs from a remote distance. Compensation pan-tilt angles \( \hat{\alpha} = (\alpha, \beta) \) for visual landmark alignment can be derived as follows:

\[ \hat{\alpha}(t) = \hat{k} \cdot \mathbf{l}(t), \ldots (7) \]

where \( \hat{k} \) denotes constant coefficient between pixel distance in image and pan tilt angle, and \( \mathbf{l} \) denotes compensation vector to desired visual landmark position \( \mathbf{d} \) in images for visual landmark alignment which can be defined as

\[ \mathbf{l}(t) = \mathbf{d} - \hat{\mathbf{e}}^0(t). \ldots (8) \]

Then, pan-tilt mirror angles of desired AOIs acquired after implementing visual landmark searching and aligning processes are stored in the PC which controls the ultrafast pan-tilt camera system for multithread viewpoint control in the object recognition stage which will be described in the following subsection.

**4.2. Multiple Virtual Cameras with Object Recognition**

The ultrafast pan-tilt camera alternatively switches its viewpoint within the AOIs acquired in Subsection 4.1 at hundreds of hertz, and the corresponding acquired images are denoted as \( I_M(k \cdot \Delta t) \) \((k = 0, 1, 2, 3, \ldots)\). The images acquired by \( j \)-th virtual camera \( I_j^V(i) \) at time \( t \) can be derived from the HFR images \( I_M \) as follows:

\[ I_j^V(i) = I_M(i), \ldots (9) \]

\[ i = (\hat{N} \cdot k + i) \cdot \Delta t, \quad (k = 0, 1, 2, \ldots) \]........... (10)

where \( \hat{N} \) denotes the thread number. Then, objects of desired category are simultaneously recognized from images captured by each virtual camera as follows:

\[ \mathcal{D}_O(I_k^V(i)) = \{ d_0^O(t), d_1^O(t), \ldots, d_P^O(t) \}, \ldots (11) \]

where \( \mathcal{D}_O \) denotes detection operation for desired object category.

Tiny YOLOv3 [49] is used as the detector for QR code \( \mathcal{P}_L \) in visual landmark searching process and person \( \mathcal{P}_O \) in multithread AI camera mode. Here, QR code and person categories are pre-trained by Dubská and COCO dataset, respectively. Here, the AOIs for object recognition are periodically updated by repeating the landmark searching process with predetermined updating period.

**4.3. Multi-GPU-Based Recognition Acceleration**

To enable real-time recognition for HFR images \( I_3(k \cdot \Delta t) \) and \( I_M(k \cdot \Delta t) \) which are captured in visual landmark searching and multi-thread object recognition process, we accelerate the recognition speed by parallely implementing multiple CNNs with multiple GPUs. Each CNN is loaded on single corresponding GPU memory in advance before executing HFR recognition. GPU allocation policy of each input image captured at \( t = k \cdot \Delta t \) is as follows:

\[ \bar{t} = k - \left[ \frac{k}{N} \right] \cdot \bar{N}, \ldots (12) \]

where \( \bar{N} \) and \( \bar{t} \) denote GPU amount and number of implemented GPU for input image captured at \( t = k \cdot \Delta t \). Execution speed is accelerated \( N \) times compared with the execution speed when implemented in single GPU. Since execution speed of the Tiny YOLOv3 is around 150 fps with single GPU, execution speed capability of our six-GPU-mounted server is around 900 fps.

**5. Experiments**

**5.1. Visual Landmark Recognizability Evaluation**

To verify the performance of our proposed system in visual landmark searching, we evaluated its recognizability with QR code patterns of different size and module number when a wide FOV area was measured by our high-resolution imaging system. The experimental setup for recognizability evaluation is depicted in Fig. 5(a). The QR code patterns for evaluation were placed 5 m in front of the ultrafast pan-tilt camera. A 100-mm telephoto lens was mounted on the camera head, and the imaging area was \( 5.0 \times 1.0 \) m. Fig. 5(b) shows the QR code patterns used for the evaluation. Seven square QR code patterns with edge lengths of 8 cm, 6 cm, 5 cm, 4 cm, 3 cm, 2 cm, and 1 cm were printed on an A4 paper. Ten A4 papers of QR code patterns with different module numbers \((21^2, 25^2, 29^2, 33^2, 37^2, 41^2, 45^2)\),

Journal of Robotics and Mechatronics Vol.34 No.5, 2022 1057
49^2, 53^2 and 57^2) were placed left-to-right in the imaging area. 15 × 4 sub-areas of imaging area that are alternately captured by 720 × 540 images at 500 fps and consist of seamless 10,800 × 2,160 ultra-high-resolution images at 8.3 fps. The QR code pattern regions detected from a captured ultra-high-resolution images are depicted in Fig. 6. We evaluate quality of the images captured by our ultra-fast pan-tilt camera system by comparing with same area images captured by a 4K resolution video camera (GZ-RY980-A, JVC KENWOOD Corp., Japan). Fig. 7 shows the QR code patterns of different size captured by our proposed system and the 4K video camera. We can see that our system is much superior to the high-end video camera in image quality since one pixel corresponded to 0.46 mm and 1.31 mm, respectively. Next, we evaluate performance of our system by investigating the recognizability of QR code patterns with different size and module number. Here, the localized QR code patterns in CNN-based detection were decoded with ZXing Library [a]; Table 1 summarize the recognizability with QR code patterns acquired by the 4K resolution video camera. From the table, we can see that recognizability falls when the module number of QR code pattern increases, and our system can recognize and decode up to 3-cm-size QR code patterns at 5-m distance, much superior to 4K-video-camera-based recognition which can only recognize 6 or 8-cm-size QR code patterns with small module number.

5.2. Person Recognition with Multi-Thread Cameras

To verify the effectiveness of the proposed system, we present the recognition results of persons in a spacious meeting room when twenty visual landmarks are pre-allocated on desks for fixing viewpoint of each virtual camera. Fig. 8 shows the experimental environment. Three rows of desks were placed 7 m, 11 m, and 15 m in front of the multi-thread AI camera system. Twenty QR code patterns with edge length of 10 cm and their corresponding number labels were pre-placed on desk of each
Multi-Thread AI Cameras Using High-Speed Active Vision System

Fig. 8. Experimental environment when observing persons in spacious conference room.

Fig. 9. Recognized visual landmarks after scanning.

Fig. 10. Pan and tilt angles of visual-landmark-detected viewpoints.

seat. A 50-mm-focal-length lens whose horizontal and vertical angles of view were 10.5° and 7.9° was mounted on the camera head.

Figure 9 shows a panoramic image and recognized QR code regions which were generated after completing a cycle of visual landmark searching process for 12 × 8 seamless viewpoints in a wide angle area of 126° × 63° with 192 ms. From this figure we can see that all the pre-placed QR codes were recognized in the visual landmark searching process. The pan and tilt angles of twenty visual-landmark-detected viewpoints after conducting alignment are depicted in Fig. 10. Here the desired visual landmark centroid position was set to \( d = (360, 480) \) in 720 × 540 image to monitor upper regions of visual landmarks. Three persons came in the meeting room and sit in randomly selected seats. Fig. 11 shows output images of twenty virtual AI cameras at a moment, where we can see the three persons were robustly recognized in images. Here, refresh rate of the output images for object recognition was 25 Hz.

To confirm quality of the acquired image by using our proposed system, we implemented face identification method ArcFace [50] on acquired images when persons sit in the seats. Fig. 12(a) shows face-identified results for 400 × 400 ROI images of persons which were cropped from 720 × 540 input images. All the persons in the camera view were registered for face identification in advance. For comparison, the face-identified result when ArcFace was implemented on 200 × 200 ROI images cropped from images captured by the still 4K camera used in Subsection 5.1, is shown in Fig 12(b). From the identified result we can see that persons captured by our system were correctly identified, while 4K-camera-captured images are too degraded for identification.
6. Conclusion

In this study, we proposed a multi-thread AI camera system that can simultaneously recognize remote objects in desired multiple AOsIs in wide FOV by using single image sensor. The system consists of an ultrafast active vision system and a CNN-based ultrafast object recognition system. The ultrafast active vision system can function as multiple virtual cameras with high spatial resolution by synchronizing exposure of a high-speed camera and movement of an ultrafast two-axis mirror device at hundreds of hertz, and the CNN-based ultrafast object recognition system simultaneously recognizes the acquired high-frame-rate images in real time. The performance of the proposed multi-thread AI camera system was demonstrated by conducting several wide area monitoring experiments on persons in nature spacious scenes such as meeting room, which was formerly too wide for a single still camera with wide angle lens to simultaneously acquire clear images. We plan to improve the recognizability of our system by introducing longer focal-length lens, faster pan-tilt camera, and higher resolution camera. We also plan to apply our system to other circumstances such as wide area surveillance, robot navigation, and structural health monitoring.

Acknowledgments

This work was supported in part by the Japan Science and Technology Agency (JST) ACEIL under Grant JPMJAC1601 under Grant 18160801.

References:


Supporting Online Materials:

### Name: Kohei Shimasaki

**Affiliation:**
Specially Appointed Assistant Professor, Digital Monozukuri (Manufacturing), Education, and Research Center, Hiroshima University

**Address:**
3-10-32 Kagamiyama, Higashi-hiroshima, Hiroshima 739-0046, Japan

**Brief Biographical History:**
2016 Received B.E. degree from Hiroshima University
2018 Received M.E. degree from Hiroshima University
2019- Specially Appointed Assistant Professor, Digital Monozukuri (Manufacturing), Education, and Research Center, Hiroshima University
2020 Received Ph.D. degree from Hiroshima University

**Main Works:**
- High-speed vision, vibration imaging

**Membership in Academic Societies:**
- The Iron and Steel Institute of Japan (ISIJ)
- Institute of Electrical and Electronics Engineers (IEEE) Robotics and Automation Society (RAS)
- The Japan Society of Mechanical Engineers (JSME)

### Name: Shaopeng Hu

**Affiliation:**
Smart Robotics Laboratory, Graduate School of Advanced Science and Engineering, Hiroshima University

**Address:**
1-4-1 Kagamiyama, Higashi-hiroshima, Hiroshima 739-8527, Japan

**Brief Biographical History:**
2013 Received B.E. degree from Northeastern University
2015 Received M.E. degree from Northeastern University
2018 Received Ph.D. degree from Hiroshima University
2018- Designated Assistant Professor, Graduate School of Advanced Science and Engineering, Hiroshima University
2020- Assistant Professor, Graduate School of Advanced Science and Engineering, Hiroshima University

**Main Works:**
- High-speed active vision, stereo measurement, deep intelligence systems

**Membership in Academic Societies:**
- Institute of Electrical and Electronics Engineers (IEEE)
- The Japan Society of Mechanical Engineers (JSME)

### Name: Idaku Ishii

**Affiliation:**
Smart Robotics Laboratory, Graduate School of Advanced Science and Engineering, Hiroshima University

**Address:**
1-4-1 Kagamiyama, Higashi-hiroshima, Hiroshima 739-8527, Japan

**Brief Biographical History:**
1992 Received B.E. degree from The University of Tokyo
1994 Received M.E. degree from The University of Tokyo
1996-2000 Assistant Professor, The University of Tokyo
2000 Received Ph.D. degree from The University of Tokyo
2000-2002 Lecturer, Tokyo University of Agriculture and Technology
2002-2003 Associate Professor, Tokyo University of Agriculture and Technology
2003-2007 Associate Professor, Hiroshima University
2007- Professor, Hiroshima University

**Main Works:**
- High-speed robot vision, dynamic sensory information processing, sensor-based robot control, and their smart applications

**Membership in Academic Societies:**
- Institute of Electrical and Electronics Engineers (IEEE)
- The Robotics Society of Japan (RSJ)
- The Japan Society of Mechanical Engineers (JSME)
- The Society of Instruments and Control Engineers (SICE)