A Study on Motion-blur-free High-frame-rate Video Camera Using High-speed Mirror Actuator

（高速ミラーアクチュエータを用いた
モーションブラーフリー高速度カメラ
に関する研究）

by

Michiaki Inoue
井上　満晶

Graduate School of Engineering
Hiroshima University
March, 2020

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Motion blur is a well-known phenomenon that occurs when shooting images of fast-moving scenes. The degradation degree of images caused by motion blur depends on the duration of the camera exposure, as well as on the apparent speed of the target scenes, and the camera's exposure time is often decreased in order to reduce motion blur. However, a trade-off exists between brightness and motion blur in image shooting, because it is difficult to obtain non-blurred bright images with a decreased exposure time since less light is then projected onto the image sensor. This trade-off is extremely aggravated in highly magnified observations of fast-moving scenes in various application fields which may not use a strong illumination, such as flowing cells in microscopic fields, precise inspections of products on a moving conveyor line, and road surface and tunnel wall inspections from a moving vehicle, because the apparent speed of the scene increases in the magnified camera view, and the light that is projected on the image sensor diminishes when the magnification is increased.

1.2 Onsite activities in TAOYAKA program

To summarize the problems of the motion-blur in the various areas, I have joined several onsite activities in the TAOYAKA program. Through these activities, I have explored actual applications related to the motion-blur, especially in disadvantaged areas.

In the first activity, I had joined an onsite education and learned the concept of

health monitoring of infrastructure in the TAOYAKA program. Recently, Japanese governments obligated the periodic inspection of the infrastructures. The required demand of the evaluation index for road tunnel inspections by the robot technology from the Japanese government in 2018 is recognitions of a crack on the tunnel wall with 0.3-mm-width [1]. As seen from this parameter, it needs high-quality images of damaged areas for health monitoring. Under these circumstances, several inspection systems from a fast-moving vehicle [2] have already been developed and have succeeded in reducing inspection time and labor costs. The inspection speed of the vehicle has a trade-off relationship between the efficiency and the precision because the high-speed scanning from the moving vehicle deteriorates the quality of images due to motion blur.

In the second activity, as a reverse innovation approach in the TAOYAKA program, I had surveyed a hilly and mountainous area in the Sera city in Hiroshima prefecture in Japan. In the hilly and mountainous area of Japan, agricultural developments are facing more and more problems such as abandoned farmland, a lack of labor force, and a lowland cultivation ratio. Under such backgrounds, it is urgent and necessary to develop comprehensive monitoring of agricultural land use, thus to assess the agricultural resource management and try to improve it. Like most cities in hilly and mountainous areas of Japan, Sera also has a depopulation and an aging population problem. Most farmers also have no successors that cause the problem of the labor shortage. Thus, I proposed agricultural supports by video surveillance as one solution for the above issues in the Onsite-Team-Project. I had clarified the farmers' need for the system specifications as the goal in this project. I surveyed two unions: Kirarikariyama (15 households) and Kurohada (9 households) to have a comprehensive understanding of their demands. As a result, the frequency to look around is once in several days in the evening or early morning. The image quality of cameras is less than averaged 0.74 pixels/mm. The farmers need high-quality images of the crop when monitoring the broad agricultural fields; therefore, a new high-resolution scanning system using a high-speed flying drone may be required. In the system, the high viewpoint velocity in images would occur. In contrast, it might need a long exposure time for shooting high-dynamic-range images when it keeps

the recognizability of crop in images. Hence, the trade-off problem of the motion-blur in this situation will exist as well. The section of the appendix summarized this activity of the Onsite-Team-Project.

As described above, a new video camera system that can capture non-blurry and bright images of fast-moving objects with the long exposure time of the camera is required for various actual applications, including the disadvantaged areas.

1.3 Overview of this study

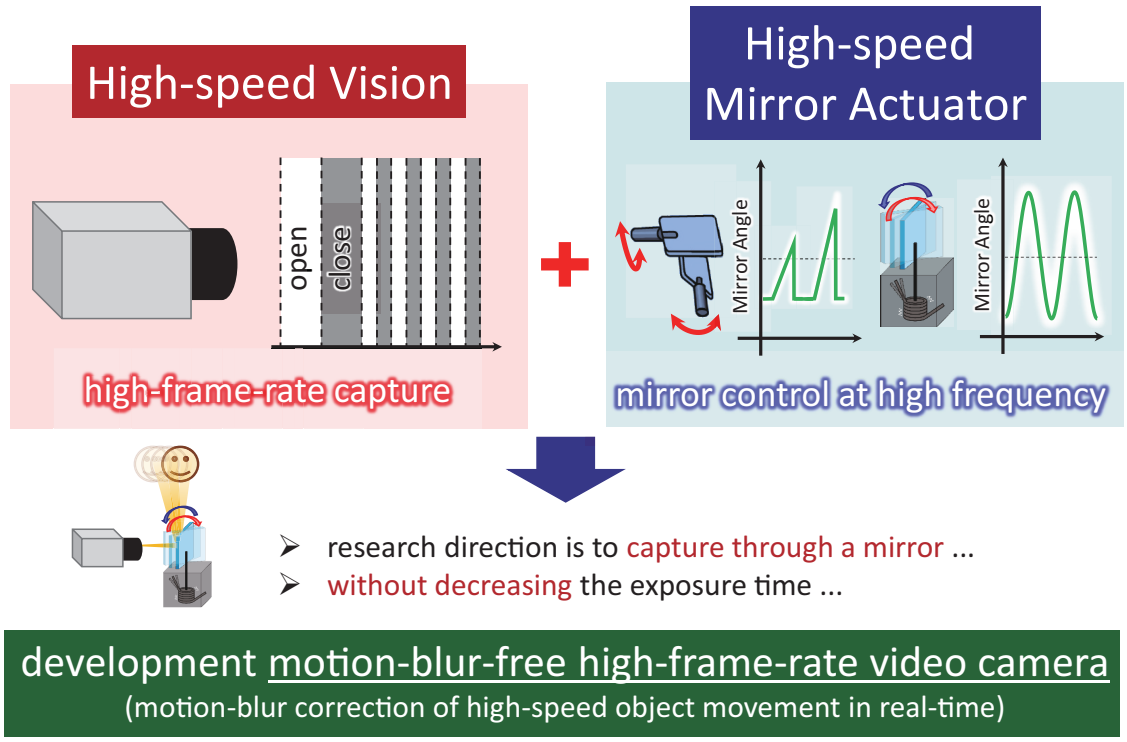


Figure 1.1: Concept overview of this study.

In this study, we develop a motion-blur-free high-frame-rate (HFR) video camera that can capture non-blurred images of fast-moving objects without lowering the camera's exposure time. It realizes motion-blur-free high-frame-rate video capturing for fast-moving objects as our research direction is that a camera captures the objects via a mirror actuator. The motion-blur-free HFR video camera consists of a high-speed vision that can capture an HFR image at variable frame-timing and a high-speed mirror actuator, which

controls its mirror angle at a high frequency. We propose novel concepts in order to realize a real-time motion-blur-free HFR video shooting for high-speed moving objects. Figure 1.1 shows the concept of our research.

The first concept is a camera-driven method with a high-speed mirror actuator for frame-by-frame intermittent tracking. Motion blur in video shooting of moving objects depends on their apparent motions on the image sensor when the shutter is open, that is, when the incident light accumulates on the image sensor, whereas their apparent motions cause no motion blur when the shutter is closed, that is, when the image sensor is blind to any incident light. Thus, in this tracking method, the mirror-based vision system shoots non-blurred and bright videos of fast-moving scenes with a fixed camera position by alternating the tracking control methods according to whether the shutter is open or closed. This concept was mentioned in Chapter 3.1.2 in detail. We developed a motion-blur-free video shooting system that simultaneously controls the angles of the pan and tilt mirrors on a 2-DOF piezo actuator-based active vision system for HFR video shooting through the implementation of a frame-by-frame intermittent tracking algorithm in real-time, and conducted several experiments in Chapter 4.

The second concept is an actuator-driven method for frame-by-frame intermittent tracking. Building on the camera-driven method with a high-speed mirror actuator in which the actuators are simultaneously controlled for tracking in synchronization with the camera's frame timings, we extend the method to the actuator-driven method so that the camera's frame-capture-event and the mirror amplitude are controlled for motion-blur-free video shooting in synchronization with the large amplitude vibration of a free-vibration-type actuator such as a resonant mirror vibrating at a high-frequency corresponding to its natural frequency. The proposed concept can derive the maximum performance of the free-vibration-type actuator that enables motion-blur-free shooting of faster-moving objects at a higher frame rate. This concept was described in Chapter 3.1.3 in detail, implemented in the prototype system of Chapter 5 and Chapter 6. Its effectiveness was verified through several experiments.

The third concept is a blur-index-based visual-feedback algorithm. In order to solve

the problem of a conventional object-position-based visual-feedback algorithm in the other motion-blur-free video camera systems, this concept realizes a visual-feedback control for a camera viewpoint-velocity to facilitate motion-blur-free high-frame-rate video recording of moving objects without relying on object patterns at a higher frequency. In this research, we utilized an edge-intensity calculation in image processing as the blur-index. The implemented edge-intensity-based visual-feedback algorithm was executed by differencing the summation of edge-intensities of two consecutive images captured at distinct viewpoint-velocities. This concept was mentioned in Chapter 3.2 in detail, implemented in the prototype system of Chapter 6, and verified its effectiveness through this study.

1.4 Outline of Thesis

This thesis is organized as 7 Chapters, including this introduction.

Chapter 2 summarized related works on motion deblurring, image stabilization, high-speed vision, and line camera with time delay integration mode.

In Chapter 3, we proposed the concept of the camera-driven method with the high-speed mirror actuator, the actuator-driven method, and the blur-index-based visual-feedback algorithm for realizing the motion-blur-free high-frame-rate video capture. A frame-capture-event control and an amplitude control in the actuator-driven approach were proposed for the viewpoint velocity control during the exposure time.

In Chapter 4, a motion-blur-free high-frame-rate video camera using a 2-DOF piezo mirror was developed to verify the effectiveness of the camera-driven frame-by-frame intermittent tracking method with the high-speed mirror actuator for shooting non-blur and bright images of fast-moving objects.

In Chapter 5, a motion-blur-free high-frame-rate video camera using a resonant mirror was proposed to verify the effectiveness of the actuator-driven frame-by-frame intermittent tracking method for capturing clear and high-brightness images of fast-moving objects.

In Chapter 6, a motion-blur-free high-frame-rate video camera with the blur-index-

based visual-feedback algorithm was proposed to capture non-blurred image sequences even when unsteady rapid relative motion occurs between shooting complexed objects and cameras. The effectiveness of the developed camera system was verified in several experiments.

In Chapter 7, the final chapter, it summarized the contributions of this study and discussed future work.

The appendix included the report of the technical creation course in the Onsite-Team-Project in the TAOYAKA program.

Chapter 2

Related Works

2.1 Motion Deblurring

In order to reduce image degradation due to motion blur when observing moving scenes, many motion deblurring methods [3, 4] have been proposed to restore the blurred images by deconvolution using the estimated blur kernels that express the degrees and distributions of motion blur in the images. Blind deconvolution methodologies were used to estimate the blur kernels from a single image using parametric models for maximum a posteriori estimation [5–7]. In addition, various types of single-image motion deblurring methods have been proposed for correct prediction of true edges using filters [8–10], multi-scale coarse-to-fine approaches [11, 12] and reduction of ill-posed image priors in the deblurred images such as normalized sparsity priors [13], color priors [14], patch priors [15], dark channel priors [16] and smoothness priors [17]. Multi-image motion deblurring methods have been used to accurately estimate the blur kernels from multiple images such as super-resolution for consecutively captured images [18–20], a high-resolution still camera with a video camera [21, 22] and image deblurring with blurred image pairs [23, 24]. Considering a camera motion model such as a perspective motion model [25] and simplified three-DOF models [26, 27], several studies have reported motion deblurring systems by estimating the camera's egomotion with gyro sensors and accelerometers [28] or the camera's geometric location [29]. The CNN-based approaches have been presented to estimate the unknown blur functions on complex Fourier coefficients [30], motion flow estimation [31], handling kernel [32] and conditional adversarial

networks [33]. Most of these motion deblurring methods dealt with image restoration of input images degraded due to motion blur, and they did not consider the acquisition of non-blurred input images. There were limitations to the extent to which the images could be improved, and it was difficult to completely eliminate motion blur in the input images when significant changes with large displacement occurred in the images.

2.2 Image Stabilization

To reduce undesirable motion resulting from shaking or jiggling of the camera, a large number of image stabilization techniques has been developed. These techniques can be categorized into: (1) optical image stabilization (OIS) and (2) digital image stabilization (DIS). The lens-shift OIS systems have been designed to shift their optical path using optomechatronic devices such as shift-mechanisms for lens barrels [34, 35], a fluidic prism [36], a three-DOF lens platform with magnetic actuation [37], pan-tilt servo systems with a high-frame-rate camera [38, 39] and a deformable mirror [40]. For small systems such as mobile phones, the sensor-shift OIS systems have been compactly designed to shift their image sensors using voice coil actuators [41–46]. Many types of multi-DOF gimbal control systems [47–51] have been also used in OIS systems in handheld shooting and drone-based aerial videography with ready-made commercial digital cameras. These OIS systems can stabilize input images for reducing motion blur resulting from camera shake by controlling the optical path with the camera's internal sensors such as gyro sensors. However, these systems are not suitable for shooting blur-free images of fast-moving scenes when the camera is fixed. This is because the internal sensors cannot detect any apparent motion in the captured images. DIS systems can stabilize input images by compensating the residual fluctuation motion using an image processing technique that estimates the local motion vectors such as block matching [52–54], bit-plane matching [55, 56], feature point matching [57–62] and optical flow estimation [63–66]. Most of these DIS systems do not need any additional mechanical or optical device, and this feature makes them suitable for low-cost electronics. However, these systems are not suitable for capturing non-blurred input images because they cannot address existing

motion blur in the captured images. Since its origination in [67], many high-speed photography methods and systems with strobe lights [68–71] have been developed. They can shoot videos of fast-moving objects without motion blur with very short strobe pulses, whereas they cannot shoot videos of fast-moving objects at distant places under daylight conditions because ambient light becomes dominant.

2.3 High-Speed Vision

In order to track fast-moving objects with visual feedback, many real-time high-speed vision systems operating at 1000 fps or more have been developed [72–75]. Various types of image processing algorithms such as optical-flow [76], camshift tracking [77], multi-object tracking [78], feature point tracking [79] and face-tracking [80] have been implemented for HFR visual tracking accelerated by field-programmable gate arrays (FPGAs) and graphic processing units (GPUs) on high-speed vision systems. The effectiveness of high-speed vision has been demonstrated in tracking applications such as robot manipulation [81–84], multicopter tracking [85, 86], honeybee [87], microscopic cell analysis [88–91] and vibration analysis [92]. The tracking performances of most of these tracking systems are limited by the time delay of dozens of frames for convergence in tracking control, because the responsive speed of the actuator is much slower than those in the accelerated video capturing and processing in high-speed vision systems. Recently, the 1-ms auto pan-tilt system [93] using galvano-mirrors with accelerated pan-tilt actuators has achieved dynamic image control for ultrafast tracking of moving objects, and such galvano-mirror-based active vision systems can function as virtual multiple tracking cameras that can observe hundreds of different views in a second [94]. By tracking an object to be observed in the center of the camera view with visual feedback, such high-speed tracking systems can reduce motion blur without decreasing their exposure time because the apparent motion of the object to be observed can be canceled in the camera view when the tracking control works correctly. However, motion-blur-free video shooting in such systems is limited to a single target object because the viewpoints cannot be freely changed for observing other objects when the target object is tracked in the camera

view.

2.4 Line Scan Cameras with Time Delay Integration Mode

For the solution of capturing bright and non-blur images without lowering the shutter speed, many line-scan cameras with time delay integration (TDI) mode have been proposed [95–97]. Moreover, a block-matching algorithm [98] and an image restoration method [99] have been implemented for the noise compensation of system vibration. Those systems can capture the images by integrating the pixel values when shooting ranges shift little by little and get the brightness and non-blurred images. However, the synchronization of the shooting timing with the object's movement completely is especially difficult because the speed information has some noises in which case it captures the high-speed objects.

Chapter 3

Concept

3.1 Frame-by-frame Intermittent Tracking Method

3.1.1 Conventional Approach

For viewpoint-free video shooting of fast-moving objects without the motion-blur, in microscopic observation with a fixed camera view, Ueno et al. [100] developed a motion-blur-free microscope using a frame-by-frame intermittent tracking method. Motion blur in video shooting of moving objects depends on their apparent motions on the image sensor when the shutter is open, that is, when the incident light accumulates on the image sensor, whereas their apparent motions cause no motion blur when the shutter is closed, that is, when the image sensor is blind to any incident light. In the conventional frame-by-frame intermittent tracking method, it can reduce motion blur in video shooting by alternating control methods in a high-speed active vision system, from vision-based mechanical tracking control to back-to-home control, according to whether the camera shutter is open or closed; the active vision system changes the optical path to the image sensor. The microscopic tracking system realized non-blurred video shootings of unidirectionally moving objects at a high frame rate using a piezo actuator-based microscopic tracking system. However, the object speed for motion blur reduction was limited to 10 mm/s or less at submillimeter-level due to the upper limit of the movable range of the 1-DOF linear piezo stage, and it can not use for motion-blur-free video-shooting of general objects fast-moving in real space, which are two-dimensionally moving at several meters per second.

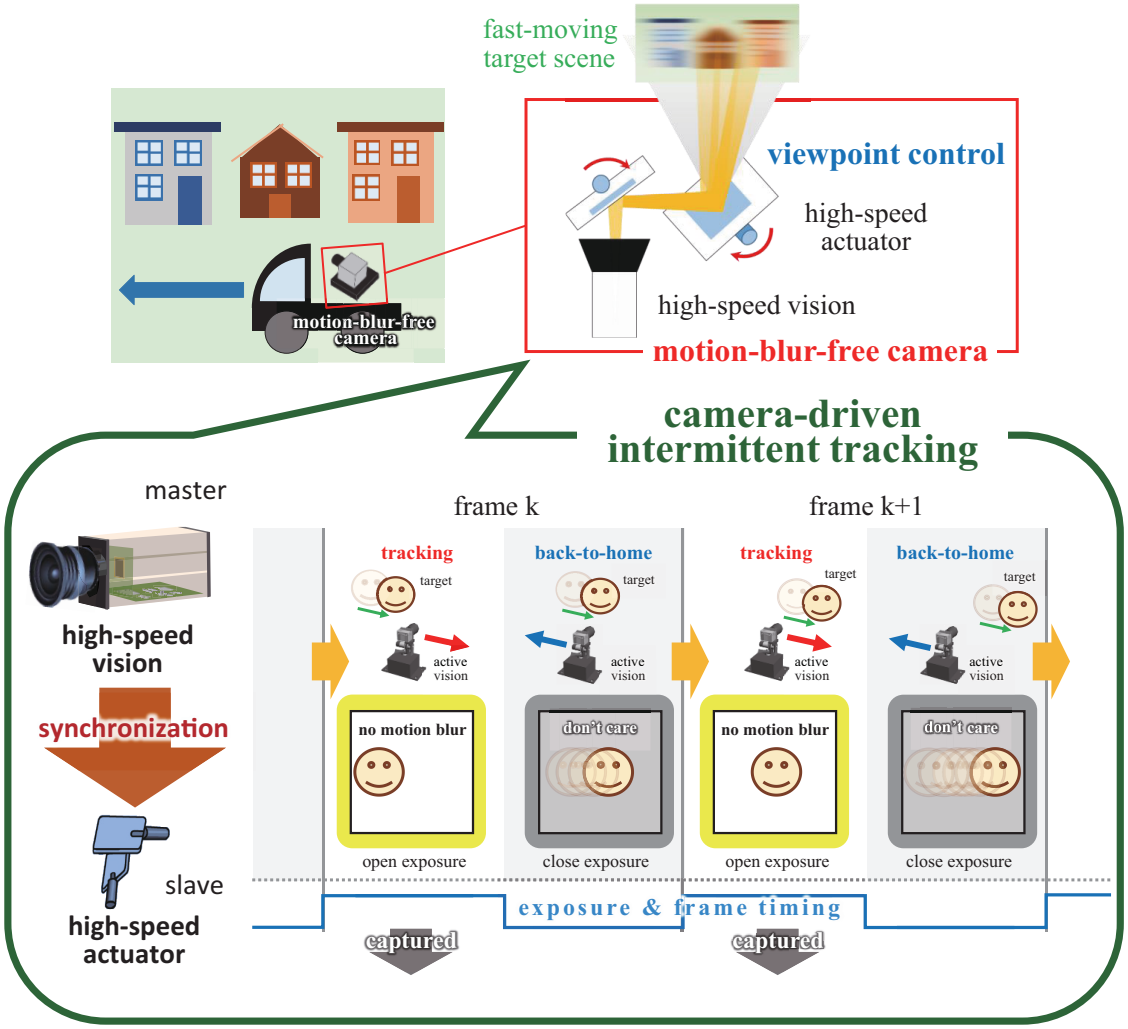


Figure 3.1: Concept of camera-driven frame-by-frame intermittent tracking with high-speed mirror actuator.

3.1.2 Camera-driven Approach with High-speed Mirror Actuator

In this study, we introduce a camera-driven frame-by-frame intermittent tracking method [100] that can reduce motion blur in video shooting by alternating control methods in a high-speed mirror-active vision system, from vision-based mirror tracking control to back-to-home control, according to whether the camera shutter is open or closed; the mirror-active vision system changes the optical path to the image sensor. This concept is illustrated in Figure 3.1. The vision-based mirror tracking control is activated to maintain the relative velocity between the coordinate systems of the object and the image sensor at zero when the shutter is open; it is operated by estimating the apparent velocity of the

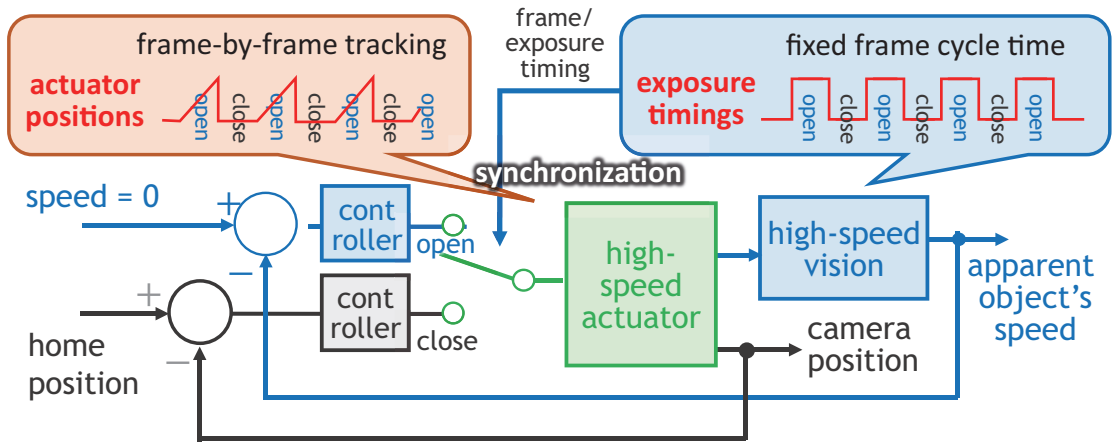


Figure 3.2: Control scheme in camera-driven frame-by-frame intermittent tracking.

objects in images in real-time. The back-to-home control is activated to reset the optical path of the camera to its home position when the shutter is closed. This control, while requiring no information from the image sensor, ensures that the movable range of the active vision system is not exceeded.

As compared with methods presented in related researches, our camera-driven frame-by-frame intermittent tracking method with the high-speed mirror actuator has the following advantages.

(1) Motion-blur-free video shooting : without decreasing the exposure time of the camera, high-brightness images of fast-moving objects can be captured without motion blur.

(2) Vision-based frame-by-frame image stabilization : without any internal sensor being required, the apparent speed of fast-moving objects on the image sensor can be controlled at zero in every frame with real-time motion estimation, which is accelerated by high-speed video processing.

(3) Free-viewpoint observation : users can freely alter the viewpoint of the camera, when it is controlled by frame-by-frame intermittent tracking. The method includes fixed-viewpoint observation.

In the camera-driven method with the high-speed mirror actuator, the frame-by-frame switching viewpoint control from vision-based tracking control to back-to-home control can be expressed with the saw-tooth-like trajectory of the position of the image

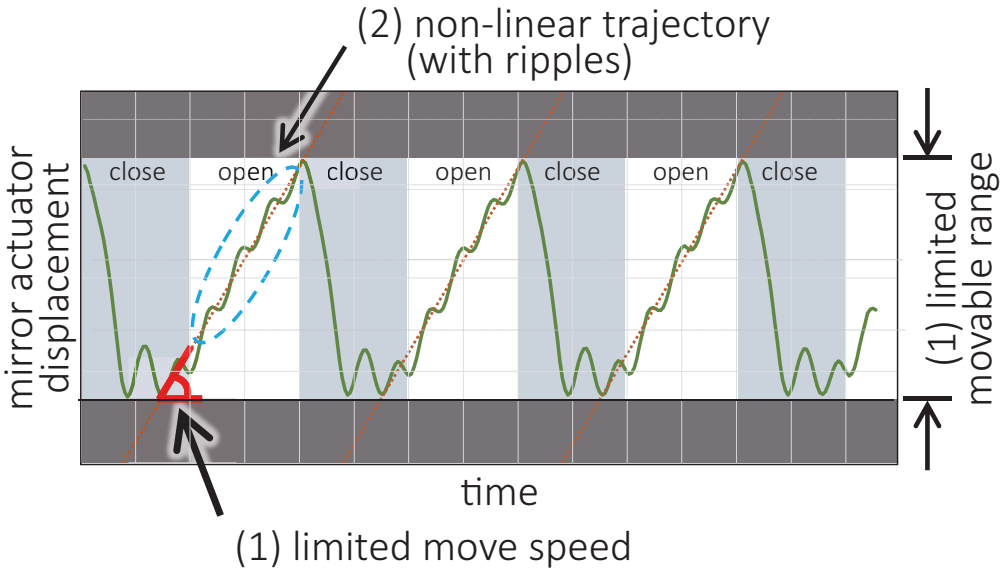


Figure 3.3: Constraints of high-frequency response actuator.

sensor desired at time t, $\boldsymbol{p}(t)$, according to whether the shutter is open or closed:

$$\boldsymbol{p}(t) = \begin{cases} \boldsymbol{p}_0 + \boldsymbol{v}(t_s(t)) \cdot (t - t_s(t)) & (0 \leq t - t_s(t) < \tau_o) \\ \boldsymbol{p}_0 & (\text{otherwise}). \end{cases}, \tag{3.1}$$

where $\boldsymbol{p}_0$ is the home position of the image sensor. τ_o and τ_c are the open and closed shutter duration, respectively. $\lfloor a \rfloor$ is the maximum integer that does not exceed a and $t_s(t) = \lfloor t/\tau \rfloor \tau$ is the time at which the image is shot at every frame; it is quantized by the frame-cycle time $\tau = \tau_o + \tau_c$. Figure 3.2 illustrates the saw-tooth-like desired trajectory and control chart of our frame-by-frame intermittent tracking method. In Eq. (3.1), the upper line expresses the control target for vision-based tracking control when the shutter is open, and the lower line expresses the control target for back-to-home control when the shutter is closed. It is assumed that the relative velocity between the coordinate systems for the object and the image-sensor at time $t_s(t)$, $\boldsymbol{v}(t_s(t))$, is estimated by processing the captured images in real-time. Thus, the state of the shutter periodically changes between open and closed in every frame, and the two control methods should switch when the video is shot with frame-by-frame intermittent tracking, corresponding to the frame-cycle time τ.

For HFR video shooting of moving objects with camera-driven frame-by-frame intermittent tracking, a high-speed actuator that can periodically drive the motion of an image sensor at hundreds of hertz or more should be accelerated for frame-by-frame switching of the control methods; its switching frequency perfectly corresponds to the frame rate of a high-speed vision system that can estimate the apparent speeds of moving objects on the image sensor in real-time. To design a motion-blur-free video shooting system with camera-driven frame-by-frame intermittent tracking of moving objects, the following constraints pertaining to a high-frequency response actuator should be considered, as well as the frame rate of the vision system, as illustrated in Figure 3.3.

(1) Limited movable range and move speed

A high-frequency response actuator has to perform a trade-off between its frequency response and movable range. The amplitude of the repetitive motion at a high frequency is limited because the movable range of a high-frequency response actuator gets narrower as its mechanical time constant gets smaller. In camera-driven frame-by-frame intermittent tracking with the camera exposure time, the high-frequency response actuator should continuously track a target object whenever the camera shutter is open. However, the motion blur cannot be completely eliminated when the distance of the object during the time the camera shutter is open is larger than the movable range of the actuator. The admissible speed of the target object is limited in motion-blur-free video shooting with a large camera exposure time.

(2) Limited controllability in the high-frequency range

A high-frequency response actuator requires a certain time to attenuate its ringing response with resonant vibration because it achieves its high-frequency drive with a low damping ratio by reducing its viscosity such as friction. The trajectory of a high-frequency response actuator should be linearly controlled whenever the camera shutter is open so as to cancel the apparent speed of the target object, assuming that it moves at a fixed speed as long as the camera shutter is open. However, it is difficult to completely eliminate ripples in the actuator's trajectory in the camera-driven frame-by-frame intermittent tracking at hundreds of hertz or more because the frame interval is not larger than its damping time

for resonant vibration, and motion blurs are still retained in the images.

3.1.3 Actuator-Driven Approach

According to the constraints stated in the previous section, the camera-driven method with the high-speed mirror actuator using piezo-mirror [101, 102] and galvano-mirror [103, 104] cannot always derive the maximum performance of a high-frequency response actuator, and the frame rate of a high-speed vision system should be lowered so as to maintain the linear trajectory of the actuator during the shutter open of the camera. The very flexible controllability of the high-speed vision system, whose frequency response is much higher than that of the actuator, was not fully utilized in the frame-by-frame intermittent tracking.

Thus, in this section, we propose an improved frame-by-frame intermittent tracking method that can reduce motion blur in video shooting by controlling the camera shutter timings in synchronization with the resonant vibration of a free-vibration-type actuator such as a resonant mirror. Its high-frequency vibration with a large amplitude enables the ultrafast gaze control to track fast-moving objects during the time the camera shutter is open. Figure 3.4 shows the concept of our proposed actuator-driven frame-by-frame intermittent tracking method, and Figure 3.5 shows the control scheme of the actuator-driven approach.

When the camera's viewpoint moves unidirectionally, the viewpoint's position $x(t)$ at time t vibrates at a cycle time of $T = 1/f_0$ on the following sinusoid trajectory,

$$x(t) = A(t) \cdot \sin \frac{2\pi}{T} t. \tag{3.2}$$

where f_0 is the resonant frequency of the free-vibration-type actuator and $A(t)$ is the amplitude of the vibration at time t for controlling the viewpoint speed as an amplitude control, assuming $x(t) = 0$ when $t = 0$. The center exposure time τ_s is controlled on the following timings:

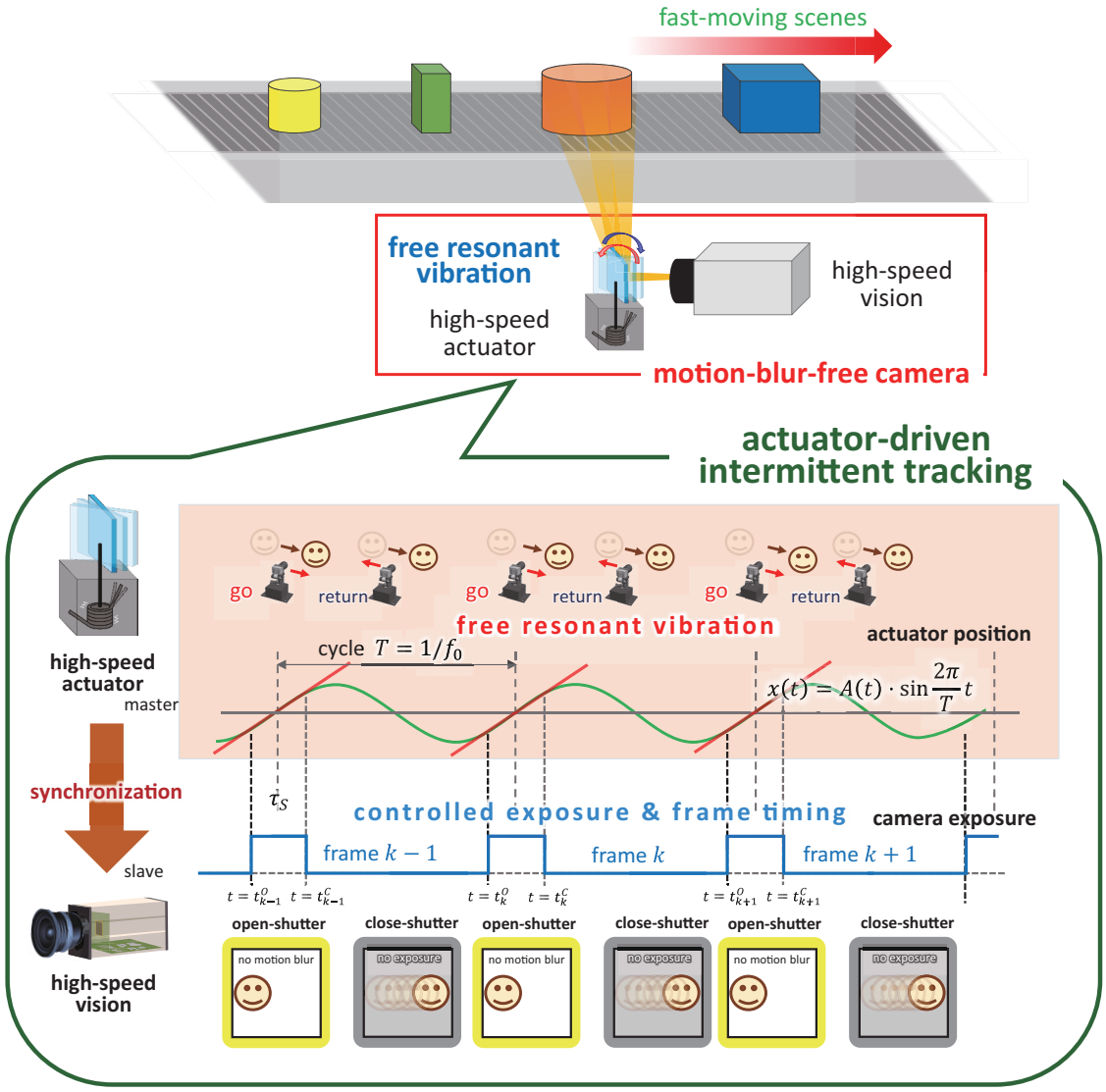


Figure 3.4: Concept of the actuator-driven frame-by-frame intermittent tracking.

$$\tau_s = t_k + \tau_d. \tag{3.3}$$

where $t_k = 2n\pi$ (n : integer) is the center time of the sinusoidal trajectory $x(t)$ in each period; and τ_d is the variable frame-capture-event that controls the slope of the approximate line to the sinusoidal trajectory when the camera shutter is open as a frame-capture-event control.

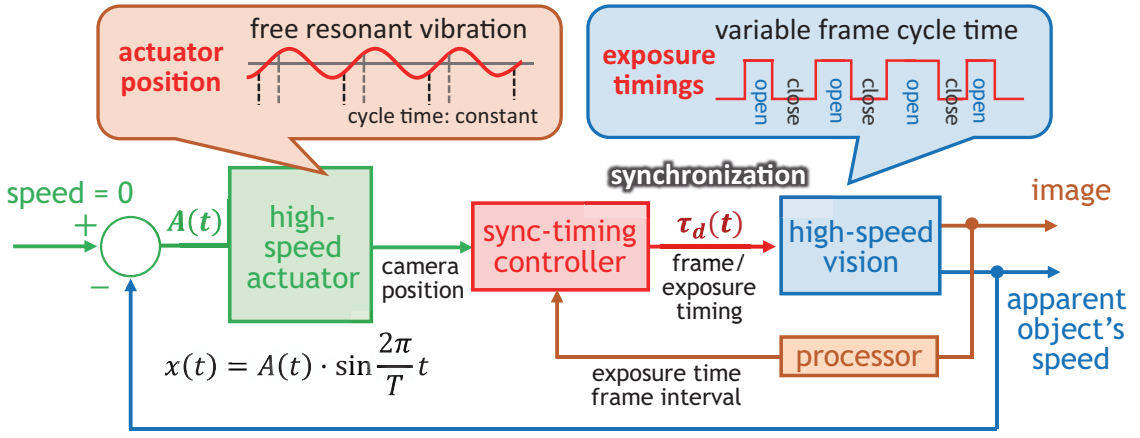


Figure 3.5: Control scheme in actuator-driven frame-by-frame intermittent tracking.

In the frame-by-frame intermittent tracking with a free-vibration-type actuator, the resonant frequency, which is a fixed value peculiar to the actuator, is not controllable, and the speed of the camera's viewpoint can be controlled with the amplitude of the vibration, as well as the exposure start and end times, which determine the time range for the linear approximation to the sinusoid trajectory. In the amplitude control of the actuator-driven tracking approach, the exposure start and end times to capture the image at frame k, which are expressed as t_k^O and t_k^C, respectively, are controlled so that the camera shutter is open when the viewpoint is located in the highly linear range within the sinusoid trajectory. The slope of the approximate line to the sinusoid trajectory when the camera shutter is open, which indicates the speed of the camera's viewpoint, is controlled for motion blur reduction so as to coincide with the apparent speed of the target object on the image sensor. In the frame-capture-event control of the actuator-driven tracking approach, the frame-capture-event of camera τ_d are controlled so that the speed of the camera's viewpoint coincides with the apparent speed of the target object within the camera exposure time, whereas the amplitude of the resonant mirror is setting as the constant value. The velocity of the camera's viewpoint is controlled with the camera shutter timings; thus, it has high responsiveness at thousands of hertz or more. Figure 3.6 shows the control method of the viewpoint velocity in our actuator-driven tracking approach.

Compared to the performance-limited mechanical actuator control in the camera-

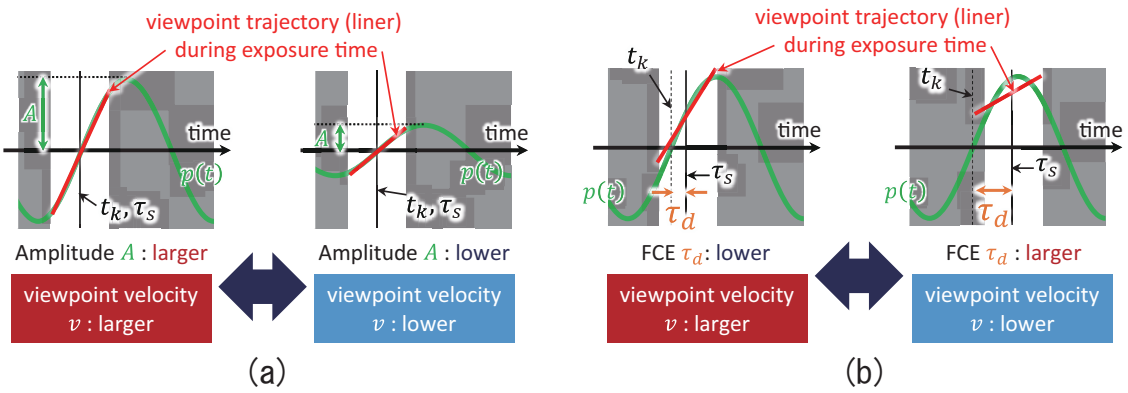


Figure 3.6: Control method of viewpoint velocity in the actuator-driven frame-by-frame intermittent tracking; (a) amplitude control and (b) frame-capture-event control.

driven tracking approach, the actuator-driven tracking approach can derive the maximum mechanical performance of a free-vibration-type actuator enables motion-blur-free video shooting of faster-moving objects at a higher frame rate, whereas a free-vibration-type actuator is plagued by the following limitations:

(1) Unresponsive amplitude control in resonant vibration

A free-vibration-type actuator tends to move on a periodic trajectory with a certain hysteresis caused by friction, and it has largely deviated from the ideal sinusoid trajectory in the case of the resonant vibration with a small amplitude. In the amplitude control of the actuator-drive tracking approach, such properties may degrade the tracking performance in video shooting a target object whose speed is either very low or varies with time.

(2) Limited time aperture ratio

In the camera-driven tracking approach, the time aperture ratio, which is the ratio of the frame interval and the exposure time in video shooting, can be programmably determined by designing the target trajectory of the camera's viewpoint freely, whereas the high-frequency response actuator cannot move on the target trajectory with a large amplitude, due to its limited movable range and speed. On the other hand, the time aperture ratio in the actuator-driven tracking approach is limited due to the sinusoid trajectory with resonant vibration. This is because the camera shutter timings are automatically determined so as to guarantee the linear motion of the camera's viewpoint when the cam-

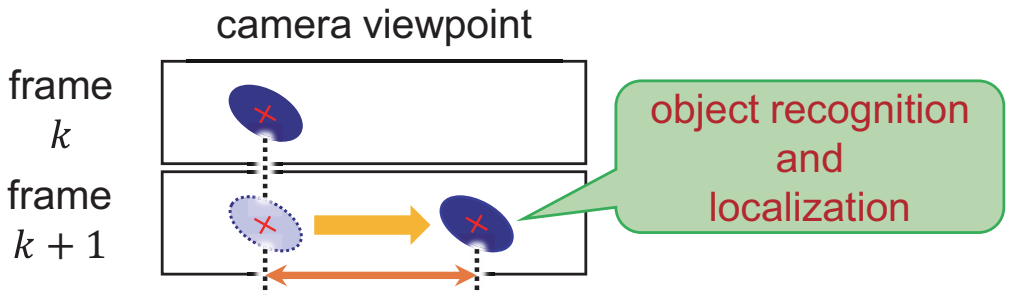


Figure 3.7: Conventional object-localization-based visual-feedback algorithm.

era shutter is open, whereas the percentage of the linear range on the sinusoid trajectory decreases as the exposure time increases.

(3) Limited delay time ratio

The delay time ratio in the actuator-driven tracking approach, which is the ratio of the mirror interval and the center time of the exposure time, is limited as well because the percentage of the linear range on the sinusoid trajectory decreases especially when the sloped of viewpoint trajectory during exposure time is small under the low-speed condition of the shooting objects. In the frame-capture-event control of the actuator-drive tracking approach, the system of tracking performance may be reduced when the shooting objects at a lower speed.

3.2 Blur-Index-Based Visual-feedback Algorithm

Conventionally the visual-feedback algorithm for viewpoint speed during exposure time has been used for motion-blur-free HFR vision systems [100–102, 104]. The object detection, recognition, and tracking approaches were efficiently used for the localization and viewpoint speed estimation. However, the robustness in the estimation rate could not be achieved due to the pattern-dependent recognition methodology. Figure 3.7 shows the conventional object-localization-based visual-feedback algorithm for the viewpoint velocity.

In this study, we proposed a blur-index-based visual-feedback algorithm method which is invariant to the patterns of shooting objects or background. It considers the

direct proportion between a blur-index of image features and the sharpness of an image. The viewpoint velocity of the vision system is controlled according to blur-index-based viewpoint speed estimation using a hill-climbing algorithm. This concept needs two blur-indexes which capture the same object are setting at slightly shifted viewpoint speeds v_1 and v_2 ($v_1 < v_2$). This method assumes maximizes the blur-index that results into clear image acquisition. The viewpoint speed relationship is calculated by subtracting slower viewpoint velocity $B(v_1)$ from faster one $B(v_2)$. The results are characterized as,

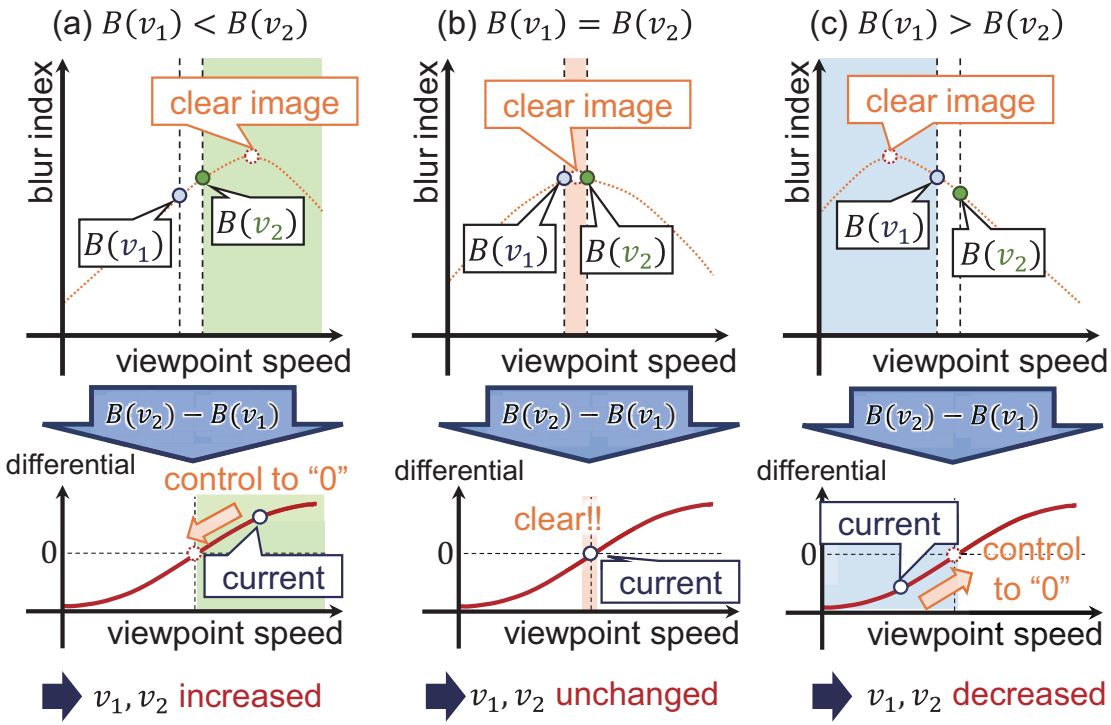


Figure 3.8: Concept of blur-index-based visual-feedback algorithm.

(a) $B(v_1) < B(v_2)$: when the difference value of blur-index (DBI) is positive that means the viewpoint velocities are slower than the actual object speed. In this case, the viewpoint velocity controller increases both v_1 and v_2.

(b) $B(v_1) = B(v_2)$: when DBI is zero that means the actual object speed is located midway in these viewpoint velocities. In this case, the controller does not change both v_1 and v_2. It indicates that the vision system has captured the clearest images which are motion-blur-free.

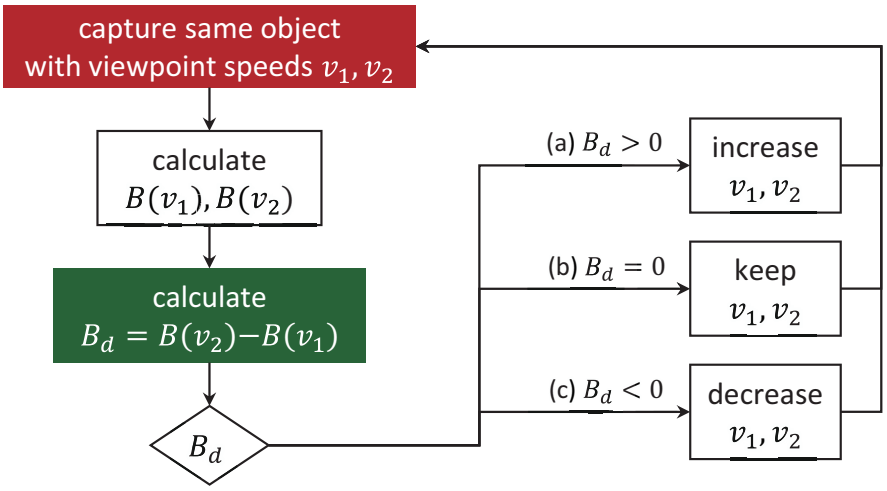


Figure 3.9: Control process of blur-index-based visual-feedback algorithm.

(c) $B(v_1) > B(v_2)$: when DBI is a negative that means viewpoint velocities are faster than the actual object speed. In this case, the controller decreases both v_1 and v_2.

Hence, we can control the camera viewpoint velocity from the frame-by-frame of any object or scene irrespective of their patterns, which is a significant advantage over conventional methods. Figure 3.8 shows the concept of the blur-index-based visual-feedback algorithm method, and Figure 3.9 shows the control process of the method.

Chapter 4

Motion-Blur-Free High-Frame-Rate Video Camera using a Piezo Mirror

In this chapter, in order to verify the effectiveness of camera-driven method with a high-speed mirror actuator, we developed a prototype motion-blur-free video-shooting system using a piezo mirror, which is designed for frame-by-frame intermittent tracking to allow zoom-in imaging of fast-moving objects without incurring motion blur.

4.1 Developed Motion-blur-free HFR Video Camera System using a Piezo Mirror

4.1.1 System Configuration

Figure 4.1 provides an overview of the prototype system when the HTZ-11000 (Joble Co., Japan) was used as the CCTV zoom lens. The system consists of a high-speed vision platform (IDP Express) [74], a CCTV zoom lens, two piezo tilt stages (PT1M36-500S-N, Nano Control Co., Japan) with mirror surfaces, and a personal computer (PC) with an ASUSTeK P6T7 WS Supercomputer mainboard, Intel Core i7 960 3.20-GHz CPU, 6-GB memory, Windows 7 Professional 32-bit OS, and a D/A board (PEX-340416, Interface Co., Japan).

IDP Express includes a camera head and an FPGA image processing board (IDP Express board). The camera head has a 512×512-pixel CMOS image sensor, the sensor and pixel size of which are 5.12×5.12 mm and 10×10 μm, respectively. The camera

Figure 4.1: Overview of motion-blur-free video camera system using a piezo mirror.

head was mounted on the camera port of the CCTV zoom lens. The IDP Express board was designed for high-speed video processing and recording, and we could implement image processing algorithms by hardware logic on the FPGA (Xilinx XC3S5000); it was mounted using a PCI-e 2.0×16 bus I/F on the PC. The 8-bit grayscale 512×512 images and processed results could be simultaneously transferred at 2,000 fps to the allocated memory in the PC.

Two piezo tilt stages were used for a mirror-drive 2-DOF active vision system to realize frame-by-frame intermittent tracking in pan and tilt directions. The piezo tilt stage can shift its surface in the rotation direction with a 2.78×10^{-6} deg resolution, and its size, weight, resonant frequency, and the range within which it can move are 36×42×29 mm, 100 g, 3,900 Hz, and 0.173 deg, respectively, when no objects are mounted on it. On the surface of the piezo stage, a 30×30×5 mm-size aluminum mirror (TFA-30S05-1, Sigma Koki Co., Japan) weighing 20 g was mounted. The piezo stage for the pan angle was installed 25 mm in front of the CCTV-zoom lens, and that for the tilt angle was installed 75 mm in front of that for the pan angle; the light from the target object passes to the tilt-mirror stage and the pan-mirror stage, and then is captured on the image sensor on the camera head. The drive voltage for the piezo stages, supplied by a high-capacity piezo driver (PH601, Nano Control Co., Japan), was 0–150 V, and the motor commands from the PC were amplified in the piezo driver in order to operate the piezo stages periodically.

In this study, a frame-by-frame intermittent-tracking algorithm was software-

implemented on the PC. The apparent speed of the objects in images was estimated in real time using the results processed on the IDP Express board, and motor commands were transferred to the piezo stage via the D/A board to reduce motion blur in the images. Corresponding to the drive voltage 0–150 V of the piezo stage, analog voltage signals in the range of 0–10.24 V were outputted from the D/A board mounted on the PC; these signals were converted at a high rate from the 12-bit digital sequences stored in the buffer of the D/A board. The details are described in the following subsection.

4.1.2 Integrated Algorithms

Assuming that a single object to be captured on video is moving two-dimensionally at a certain velocity on a plane that is parallel to the image sensor's plane, and the object's apparent velocity in images is proportional to its actual velocity on the plane, the following algorithm was implemented in the prototype system in order to observe a single object in an image.

(1) Binarization

A grayscale 512×512 input image $I(x, y, t)$ is captured at time $t = k\tau$ at an interval τ with an exposure time τ_o. $I(x, y, t)$ is binarized with the threshold I_B into $B(x, y, t)$. The apparent velocity of the object in the image at time t is estimated as $\boldsymbol{v}(t) = (\boldsymbol{c}(t)-\boldsymbol{c}(t-\tau))/\tau$ with the image centroids $\boldsymbol{c}(t) = (M_{10}/M_{00},\ M_{01}/M_{00})$ at time t and $t - \tau$. The apparent angular velocity of the object in the pan and tilt directions of a zooming optical system, $\boldsymbol{\omega}(t) = (\omega_\phi(t), \omega_\psi(t))$, is proportional to $\boldsymbol{v}(t)$ as

$$\boldsymbol{\omega}(t) = A(\boldsymbol{c}(t) - \boldsymbol{c}(t - \tau))/\tau, \tag{4.1}$$

where A is a constant parameter determined by the magnification ratio of the zooming optical system, the pixel pitch of the image sensor, and the distance between the object and the optic center of the optical system. M_{00}, M_{10}, and M_{01} are the zero- and first-order

moment features of $B(x, y, t)$ defined as

$$M_{mn}(t) = \sum_{X,Y} x^m y^n \cdot B(x, y, t), \quad (m, n) = (0, 0), (1, 0), (0, 1). \tag{4.2}$$

(2) Trajectory generation for intermittent tracking

The desired angular trajectory in the pan and tilt directions of the mirror-drive 2-DOF active vision system, $\boldsymbol{\theta}_d(t) = (\phi_d(t), \psi_d(t))$, is generated using the apparent angular velocity of the target object, $\boldsymbol{\omega}((k-1)\tau)$, which is estimated at $t = (k-1)\tau$, in order to cancel its apparent motion on the image sensor when the shutter is open from $t = k\tau$ to $k\tau + \tau_o$:

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_d(t) &= (\hat{\phi}_d(t), \hat{\psi}_d(t)), \\
&= \begin{cases} \boldsymbol{\omega}((k-1)\tau)(t - k\tau - \tau_r) + \boldsymbol{\theta}_0 & (-\tau_r \le t - k\tau < \tau_o) \\ (f(t; \phi_d(k\tau + \tau_o), \phi_0) f(t; \psi_d(k\tau + \tau_o), \psi_0)) & (\tau_o \le t - k\tau < \tau_b + \tau_o) \\ \boldsymbol{\theta}_0 & (\text{otherwise}) \end{cases} \qquad (4.3) \\
\boldsymbol{\theta}_d(t) &= \left(\text{med}(\phi_{min}, \hat{\phi}_d(t), \phi_{max}), \text{med}(\psi_{min}, \hat{\psi}_d(t), \psi_{max})\right). \qquad (4.4)
\end{aligned}$$

where $\text{med}(a, b, c)$ indicates the median value of a, b, and c. $\boldsymbol{\theta}_0 = (\phi_0, \psi_0)$ indicates the pan and tilt angles for the home position, and $[\phi_{min}, \phi_{max}]$ and $[\psi_{min}, \psi_{max}]$ indicate the movable ranges of the pan and tilt angles of the mirror-drive 2-DOF active vision system, respectively. τ_t and τ_b refer to the duration times of the vision-based tracking control and back-to-home control, respectively. $\tau_r = \tau_t - \tau_o$ is the delay time required for the mirror-drive 2-DOF active vision system to match the apparent motion of the object in the image, corresponding to its rise time. $f(t; \phi_d(k\tau + \tau_o), \phi_0)$ and $f(t; \psi_d(k\tau + \tau_o), \psi_0)$ are the fifth-order polynomial trajectory functions that ensure that the back-to-home control moves the mirror-drive 2-DOF active vision system smoothly from $\boldsymbol{\theta}_d(k\tau + \tau_o)$ to $\boldsymbol{\theta}_0$ to avoid a large acceleration. Figure 4.2 shows the timing chart for the generation of the trajectory.

(3) Control of mirror-drive 2-DOF active vision system

After storing the desired trajectory for the duration time $\tau_t + \tau_b$, the D/A board

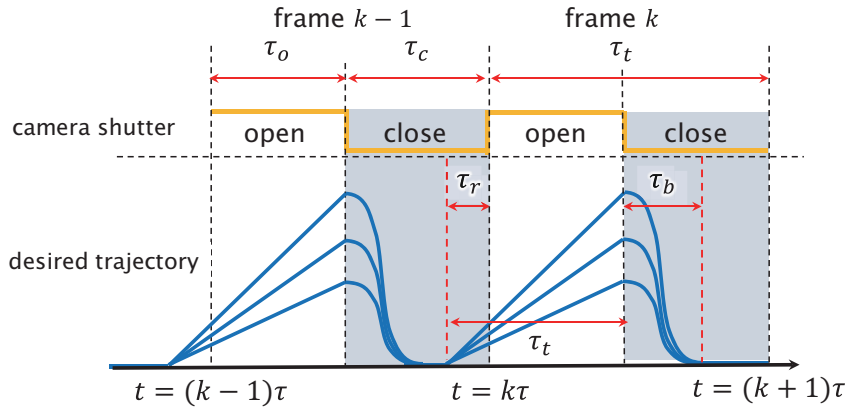


Figure 4.2: Frame-by-frame intermittent tracking trajectory.

begins sending motor commands at $t = k\tau - \tau_r$ in synchronization with the image capture timing. The motor commands for the pan and tilt angles are amplified to the drive voltages in the mirror-drive 2-DOF active vision system as

$$\boldsymbol{V}(t) = (V_\phi(t), V_\psi(t)) = (b_\phi \phi_d(t) + c_\phi, b_\psi \psi_d(t) + c_\psi), \tag{4.5}$$

where the parameters (b_ϕ, c_ϕ) and (b_ψ, c_ψ) are determined by verifying the actual pan and tilt trajectories of the mirror-drive 2-DOF active vision system.

4.1.3 Specifications

Using this prototype system, 512×512 input images were captured at 125 fps, with a frame interval of $\tau = 8$ ms and exposure time of $\tau_o = 4$ ms. The duration times of the vision-based tracking control and back-to-home control were set to $\tau_t = 4.5$ ms and $\tau_b = 2.0$ ms, respectively; the tracking time τ_t included a delay time of $\tau_r = 0.5$ ms. The desired trajectory of the pan and tilt angles of the mirror-drive 2-DOF active vision system was generated as two 16-bit digital sequences at 200 kHz for a duration of $\tau_t + \tau_b = 6.5$ ms; 1,310 16-bit data for each angle were updated with an 8-ms cycle time. The home positions of the pan and tilt angles of the mirror-drive 2-DOF active vision system were set to one end of their movable range such that $\boldsymbol{\theta}_0 = \mathbf{0}$, where their drive voltages

were 0 V. Because of the narrow movable ranges of the pan and tilt angles of the 2-DOF active vision system, 0.17 deg and 0.14 deg, respectively, the maximum speed of objects under observation without motion blur being incurred was determined theoretically by the ratio of the movable range of the duration time of the open exposure. Considering that the variations in the view angles via the mirrors correspond to twice those of the mirror angles, the maximum angular speeds for the pan and tilt angles are 67.1 deg/s and 49.7 deg/s, respectively. When the focal length of the zoom lens and the pixel pitch of the image sensor are f [mm] and $\Delta x = 0.01$ mm, respectively, one pixel corresponds to $57.3 \tan^{-1}(\Delta x/f) \approx 0.573 f^{-1}$ deg, assuming $f \gg \Delta x$; one degree corresponds to $1.75f$ pixel. When $f = 112.5$ mm, the maximum apparent speeds in the x and y directions on the image sensor for objects under observation without motion blur being incurred are 13.0 pixel/ms and 9.7 pixel/ms, respectively, corresponding to the displacements of 52.2 pixels and 38.6 pixels in the x and y directions during an exposure time of 4 ms. When $f = 650$ mm, the maximum apparent speeds in the x and y directions on the image sensor for objects under observation without motion blur being incurred are 76.1 pixel/ms and 56.4 pixel/ms, respectively, corresponding to the displacements of 304 pixels and 225 pixels in the x and y directions during an exposure time of 4 ms.

The binarization in Step 1 and the calculation of the moment features in Step 2 were implemented with parallel hardware logic for 8-bit gray-level images on the user-specific FPGA of the IDP Express board. The other steps were software-implemented as multithreaded processes with parallel executions on the PC. The execution time for Steps 1 and 2 was 0.108 ms and for Steps 3 and 4 was 0.887 ms. The total execution time was 1.01 ms. We confirmed that all the processes could be executed for 512×512 images in real time at 125 fps with an exposure time of 4 ms.

4.2 Experiments

4.2.1 Preliminary Trajectory Evaluation

First, we conducted a preliminary experiment to verify the relationship between the input voltages to the piezo stages of the mirror-drive 2-DOF active vision system and its

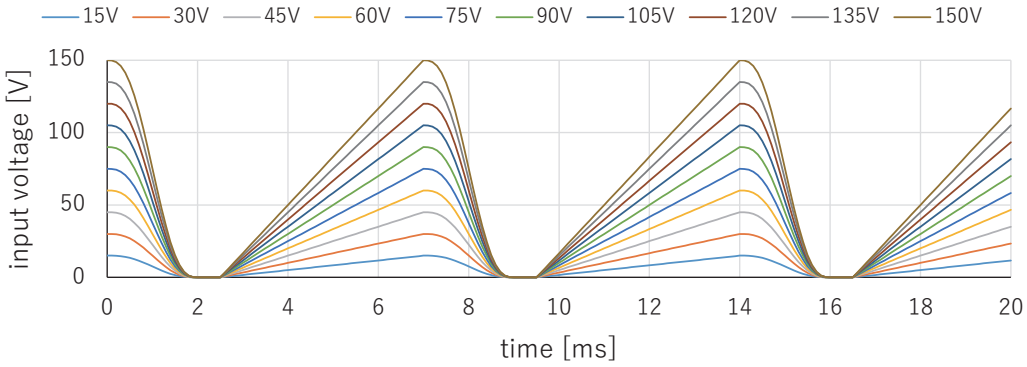
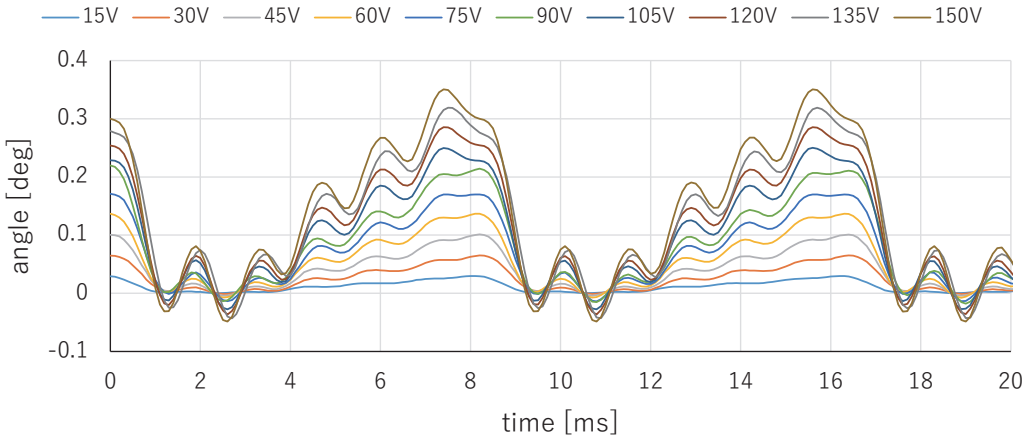
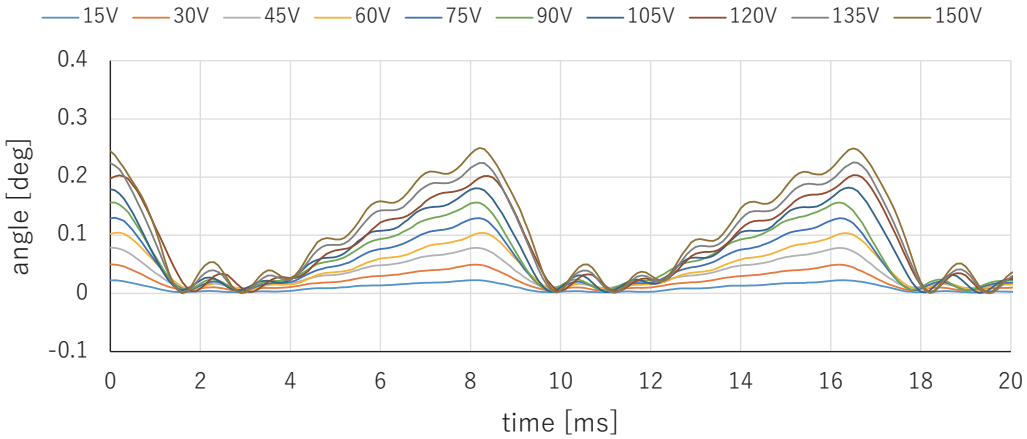


Figure 4.3: Input waveform for piezo stages of mirror-drive 2-DOF active vision system.



(a) pan



(b) tilt

Figure 4.4: Angular displacements of mirror-drive 2-DOF active vision system.

angular displacements in the pan and tilt directions when the active vision system was periodically operated on a designed trajectory at a frequency of 125 Hz. We determined

the parameters (b_ϕ, c_ϕ) and (b_ψ, c_ψ), which are expressed in Eq. (4.5), of the trajectory of the active vision system during the time when the shutter is open, and quantified the nonlinear deviations with ripples in the pan and tilt trajectories. In the experiment, the periodic voltage wave at a cycle time of $\tau = 8$ ms was inputted to the piezo stage, as shown in Figure 4.3; the input voltage wave was set to a linear waveform from 0 V to a maximum voltage V_{max} in a period $\tau_t = 4.5$ ms, where V_{max} was set to 15, 30, 45, 60, 75, 90, 105, 120, 135, and 150 V. To measure the pan and tilt angles of the active vision system, a laser beam spot for observation was redirected by the mirrors of the active vision system, and the locations of the laser beam spot projected on a screen at a distance of 4,350 mm from the active vision system were extracted offline by capturing an HFR video at 10,000 fps.

Figure 4.4 shows the angular displacements of the pan and tilt angles of the active vision system for 30 ms when the periodic input voltage waves, the maximum voltages of which varied from 0 to 150 V, at 125 Hz were applied to the piezo stages. In both the pan and tilt angles, the angular displacements were periodically changed at a frequency of 125 Hz in proportion to the amplitudes of the input voltage waves, whereas they involved certain ripple waves because of their resonant vibrations. The observed resonant frequencies in the pan and tilt angles were approximately 730 Hz and 850 Hz, respectively; they were one-fifth or less of 3,900 Hz, which is the resonant frequency of the piezo stage when no object is mounted on it. The decrease in the resonant frequencies was caused mainly the mirror attached to the piezo stage. It can be observed that the resonant frequency in the tilt angle was less than that in the pan angle, and the amplitude of the ripple in the tilt angle was more than that in the pan angle, because the tilt angular motion was more strongly affected by gravity than the pan angular motion.

When the angular trajectories during the exposure time $\tau = 4$ ms were linearized by the least squares method, Figure 4.5 shows the relationship between their inclinations and the input voltages to the piezo stages. It can be observed that the inclinations of the angular trajectories, which correspond to the apparent angular velocity of the target object, varied linearly with the amplitudes of the input voltages; the parameters in

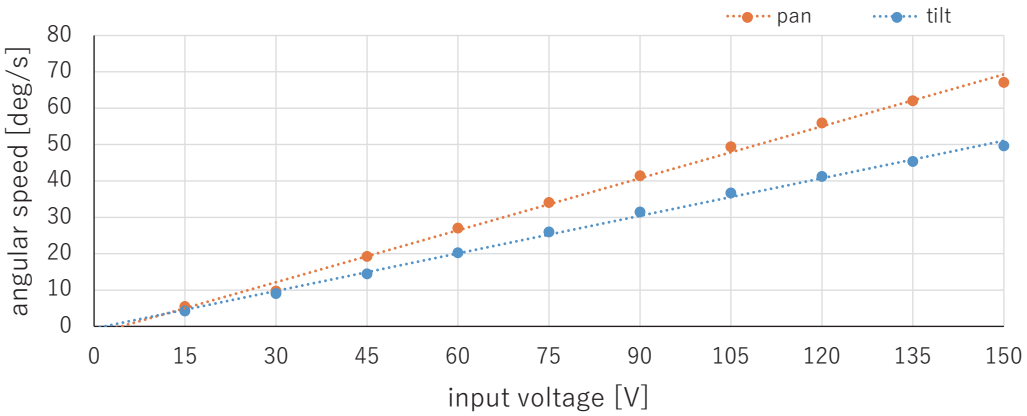


Figure 4.5: Estimated inclinations of angular displacements and input voltages.

Eq. (4.5) were estimated as $(b_\phi, c_\phi) = (2.10, 4.47)$ and $(b_\psi, c_\psi) = (2.91, 1.61)$ for the pan and tilt angles, respectively. Figure 4.6 shows the relationship between the estimated angular speeds $\tilde{\omega} = (\tilde{\omega}_\phi, \tilde{\omega}_\psi)$ and the averaged deviations $(\Delta\phi_d, \Delta\psi_d)$ from the approximated lines during 4 ms. In the figure, the ratio of the averaged deviation $(\Delta\phi_d, \Delta\psi_d)$ to the estimated angular displacement $(\phi_{mv}, \psi_{mv}) = (\tilde{\omega}_\phi\tau, \tilde{\omega}_\psi\tau)$ during the exposure time $\tau = 4$ ms, $(\Delta\phi_d/\phi_{mv}, \Delta\psi_d/\psi_{mv})$, is also plotted; the ratio indicates the percentage by which our frame-by-frame intermittent tracking method can reduce motion blur in shooting fast-moving objects. When the maximum voltage of the input image was 150 V, the angular speeds and the averaged deviations are 49.7 deg/s and 1.91 $\times 10^{-2}$ deg for the pan angle and 67.1 deg/s and 3.64 $\times 10^{-2}$ deg for the tilt angle; the ratios $(\Delta\phi_d/\phi_{mv}, \Delta\psi_d/\psi_{mv})$ were 9.3 % and 12.8 %. It can be observed that the deviation error from the approximate line becomes larger as the angular speeds become larger in both the pan and tile angles. The ratio $(\Delta\phi_d/\phi_{mv}, \Delta\psi_d/\psi_{mv})$ was not so significantly changed with the estimated angular speeds, whereas the ratio of the tilt angle was larger than that of the pan angle because of the effect of gravity. Thus, we should consider image degradation with a certain motion blur with the above-mentioned ripple deviations in motion-blur-free video shooting at 125 fps.

4.2.2 Circle-dot Motion at Constant Speeds

Next, we conducted video shooting experiments for a circle-dot pattern to verify the

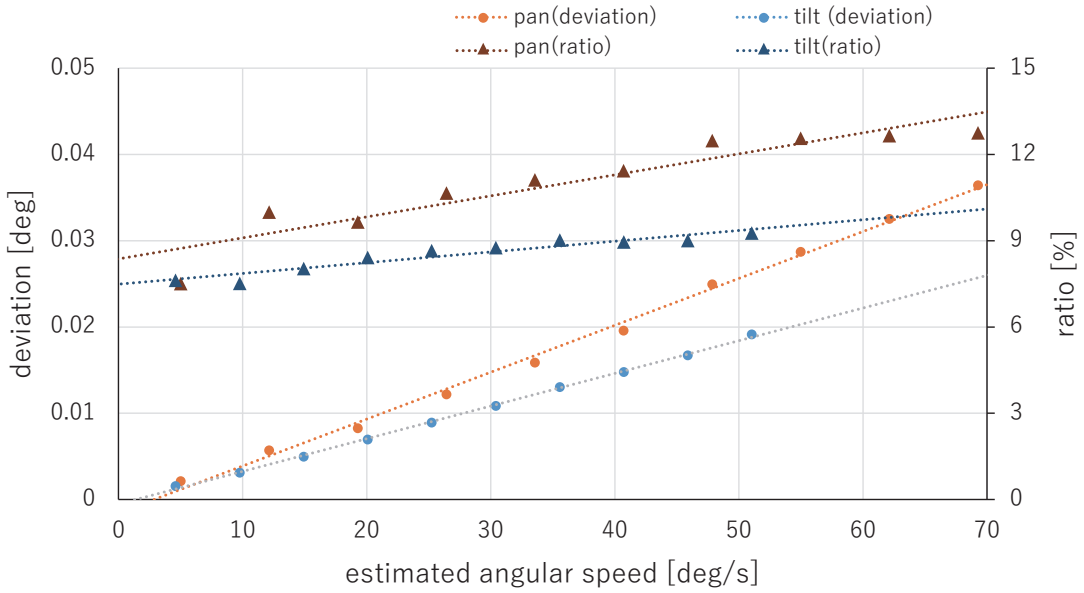


Figure 4.6: Relationship between estimated angular speeds and deviation errors.

relationship between the speed of an object and its motion blur. The pattern was moved along (1) the horizontal direction and (2) the oblique direction with an inclination of 20 deg, at constant speeds of 0, 250, 500, 750, and 1,000 mm/s using a 1-DOF linear slider. In the experiment, the HTZ-11000 (Joble Co., Japan) was used as the CCTV zoom lens; its focal length was set to f = 650 mm. The linear slider was located at a distance of 4,350 mm from the mirror-drive 2-DOF active vision system; the 35×35 mm area on a plane at a distance of 4,350 mm corresponded to an image region of 512×512 pixels, and 6.84×10^{-2} mm corresponded to one pixel. We can cancel motion blur during the 4 ms exposure when shooting a target object moving at 5.21 and 3.86 m/s on a plane 4,350 mm in front of the mirror-drive 2-DOF active vision system in the vertical and horizontal direction, respectively, corresponding to its apparent motions at 304 and 225 pixels during the 4 ms exposure time in the x and y direction on the image sensor. Figure 4.7 shows (a) an overview of the experimental environment, (b) the circle-dot pattern to be observed, and (c) the configuration of the experimental setting. The 4-mm-diameter circle dots were black-printed at intervals of 50 mm on a white sheet of paper.

Figure 4.8 shows the 227×227 images cropped from the 512×512 input images so that the circle dot is located at their centers, and Figure 4.9 shows the brightness profiles

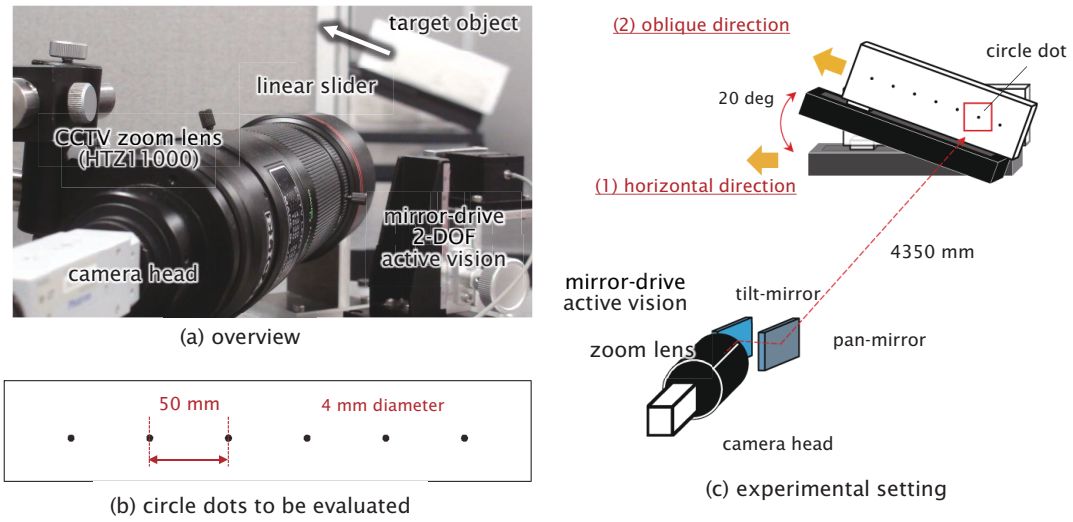


Figure 4.7: Experimental environment and circle-dot pattern to be evaluated.

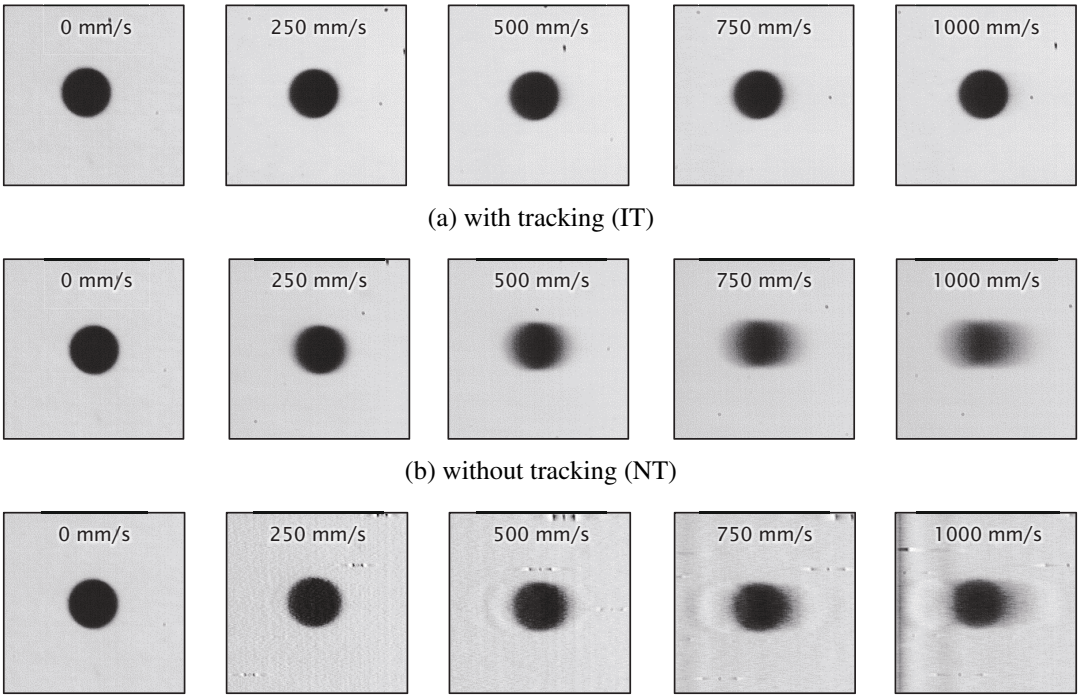
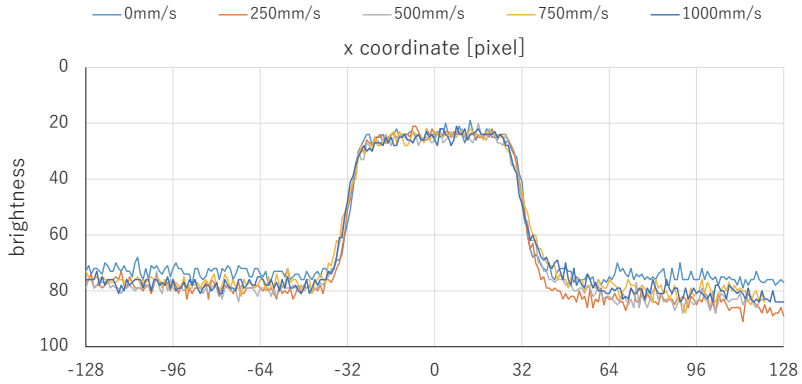
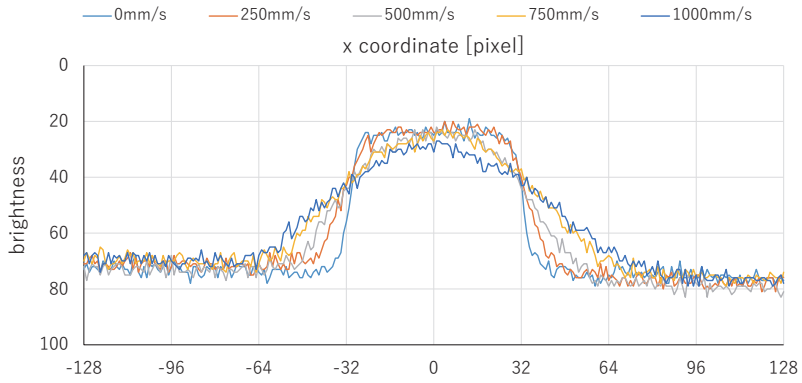


Figure 4.8: Images captured when a circle dot moved in the horizontal direction.

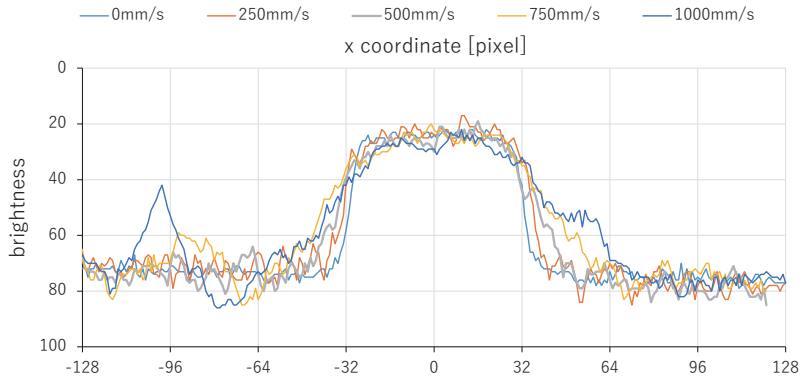
of 256 pixels on a horizontally intersected line of images when the circle dot moved at 0, 250, 500, 750, and 1000 mm/s in the horizontal direction. The threshold for binarization was $I_B = 50$. As observed in Figures 4.8 and 4.9, the input images captured with frame-by-frame intermittent tracking (IT) were compared with those captured without



(a) with tracking (IT)



(b) without tracking (NT)



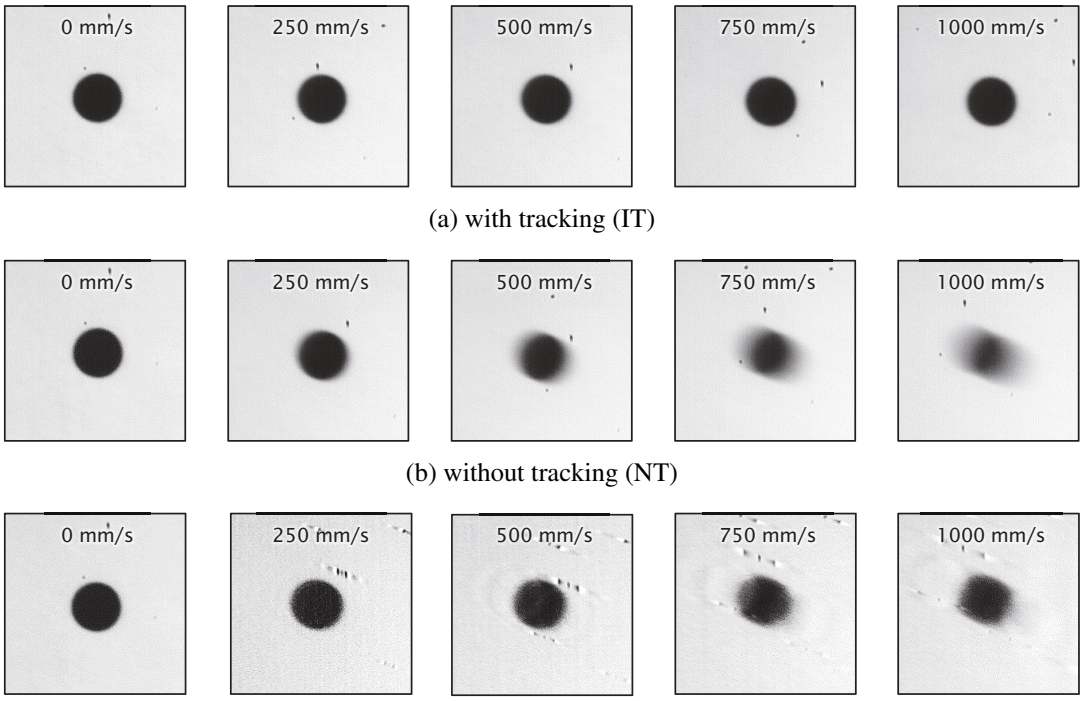
(c) motion deblurring (MD)

Figure 4.9: Intersected brightness profiles when a circle-dot moved in the horizontal direction.

mechanical tracking (NT) and their motion deblurring (MD) images. The MD images were obtained by processing the NT images offline using a non-blind convolution method with a line kernel function [110]. The NT images became increasingly blurred in the horizontal direction as the speed of the circle dot increased, whereas the IT images re-

mained almost entirely free of blurring regardless of the speed. Figure 4.10 shows the 227×227 images cropped from the 512×512 input images when the circle dot moved in the oblique direction. It can be seen that frame-by-frame intermittent tracking achieves motion-blur-free video shooting of the object moving in the oblique direction, as well as in the horizontal direction; the NT images are blurred in the 20-deg oblique direction, whereas the IT images are without blur at all the slider speeds. In the MD images, most of motion blurs were remarkably reduced, whereas certain ghost errors remained in the moving directions especially when the circle-dot moved by dozens of pixels during the camera shutter was open. This is because it is difficult for deconvolution-based methods to completely reduce large motion blurs for nonlinear brightness images with zero or saturation.

To evaluate the degree of motion blur of the observed the circle dot, the index $\Delta\lambda = \lambda_+ - \lambda_-$ was introduced; λ_+ and λ_- represent the lengths of the major and minor axes of the approximated ellipse of the circle dot in the image. The index $\Delta\lambda$ increases as the motion blur increases in the image, and is zero when the dot is a perfect circle in the image. λ_+ and λ_- were estimated offline by calculating the zero-, first-, and second-order moment features for the circle-dot region in the image, which was extracted by binarization with a threshold of 63. Considering the offset $\Delta\lambda_0 = 2.6$ pixel when no motion is present, the blur index $\Delta\lambda' = \Delta\lambda - \Delta\lambda_0$ was evaluated for the IT and NT images in Figures 4.8 and 4.10. Figure 4.11 shows the relationship between the speed of a circle dot and its blur index $\Delta\lambda'$ for the IT and NT images; $\Delta\lambda'$ was averaged for the speeds of 50 selected images. The blur index $\Delta\lambda'$ for the IT images was remarkably low at all the speeds as compared with that for the NT images; it became larger as the speed of the circle dot increased. When the circle dot moved in the horizontal direction, the blur index $\Delta\lambda'$ for the IT images was 0.9, 2.3, 3.1, and 2.7 pixel at 250, 500, 750, and 1000 mm/s, respectively; this corresponds to 13.3, 14.3, 11.5, and 7.4% of the respective value of $\Delta\lambda'$ for the NT images. When the circle-dot moved in the oblique direction, $\Delta\lambda'$ for the IT images was 0.1, 1.9, 2.6, and 2.3 pixel at 250, 500, 750, and 1000 mm/s, respectively; this corresponds to 1.6, 14.9, 12.2, and 7.5% of the respective value of $\Delta\lambda'$



(a) with tracking (IT)

(b) without tracking (NT)

(c) motion deblurring (MD)

Figure 4.10: Images captured when the circle dot moved in the oblique direction.

for the NT images. In the experiment, the speed of the circle dot was 1 m/s or less, which is considerably lower than the maximum motion-blur-free speeds of 5.21 m/s in the horizontal direction and 3.86 m/s in the and our frame-by-frame intermittent tracking method noticeably reduced motion blur of circle dots moving at all the speeds in video shooting with the exposure time of 4 ms, whereas slight motion blur remained in the IT images because of nonlinear deviations with ripples on the trajectory of the mirror-drive 2-DOF active vision system.

4.2.3 Table Tennis Ball Motion at Constant Speeds

Next, we conducted video shooting experiments for fast-moving table tennis balls launched by a table tennis machine to verify motion blur when the speed of the object to be observed is larger than the maximum motion-blur-free speed of our mirror-drive 2-DOF active vision system. Figure 4.12 shows (a) an overview of the experimental environment, and (b) the 40-mm-diameter table tennis balls that were observed. The table tennis machine (TSP Hyper S-2, Yamato Takkyu Co., Japan) was installed 4,350 mm in

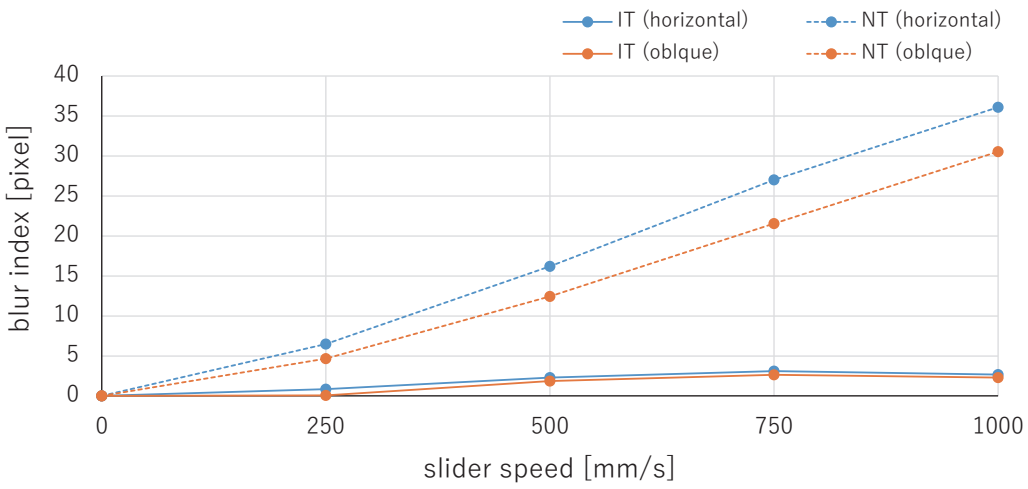
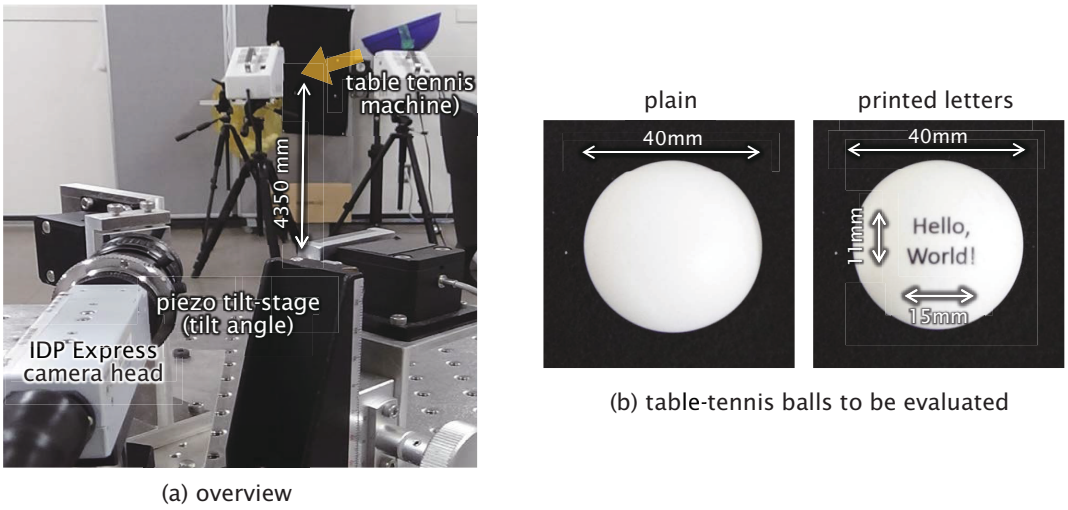


Figure 4.11: Motion blur indexes for circle dots moving at different speeds.



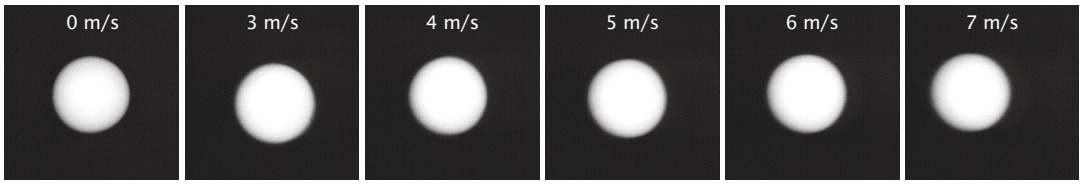
(a) overview

(b) table-tennis balls to be evaluated

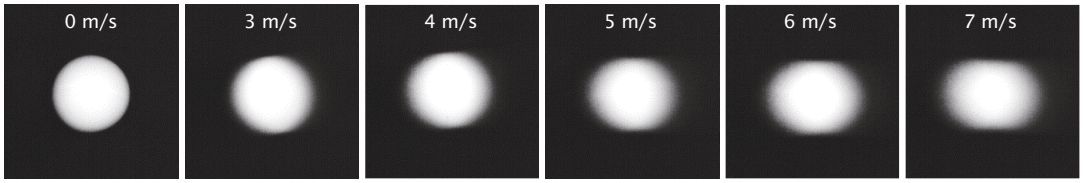
Figure 4.12: Experimental environment and observed table tennis balls.

front of the mirror-drive 2-DOF active vision system, and a table tennis ball (plain) was launched in (1) the horizontal direction, and (2) the oblique direction with an inclination of 20 deg at constant speeds of 3, 4, 5, 6, and 7 m/s. In the experiment, a CCTV lens of f = 75 mm was used with a 1.5× extender; the 200×200 mm area on a plane at a distance of 4,350 mm corresponded to an image region of 512×512 pixels and 0.391 mm corresponded to one pixel. When observing an object moving fast on a plane of 4,350 mm in front of the mirror-drive 2-DOF active vision system, the maximum motion-blur-free speeds were 5.21 m/s in the horizontal direction and 3.86 m/s in the vertical direction, corresponding to its apparent motions at 52.2 and 38.6 pixels during the exposure time of

4 ms.

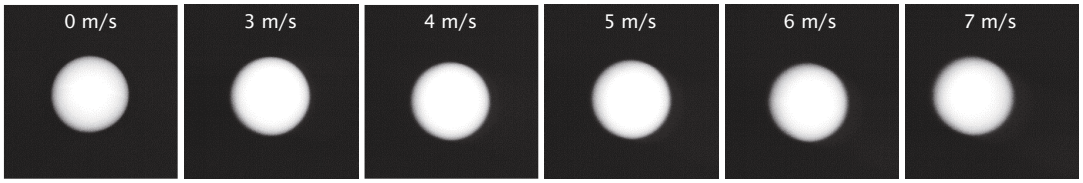


(a) with tracking (IT)

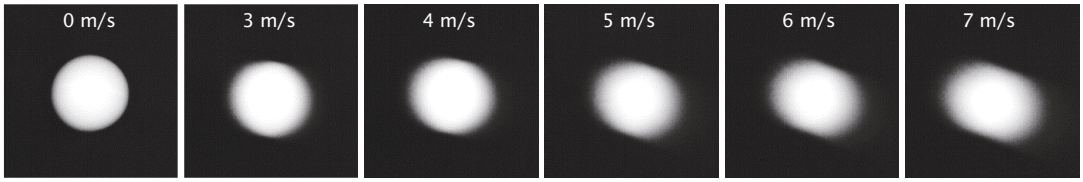


(b) without tracking (NT)

Figure 4.13: Images captured when the table tennis ball was thrown in the horizontal direction.



(a) with tracking (IT)



(b) without tracking (NT)

Figure 4.14: Images captured when the table tennis ball was thrown in the oblique direction.

Figures 4.13 and 4.14 show the 227×227 images cropped from the 512×512 input images ((a) IT images, (b) NT images) so that the table tennis ball is located at their centers when it is thrown in the horizontal direction and oblique direction. As compared with the input images captured when a table tennis ball was thrown at 3, 4, 5, 6, and 7 m/s, the input image of a motionless table tennis ball (0 m/s) is illustrated. The threshold for binarization in frame-by-frame intermittent tracking was $I_B = 50$. It can be seen that the IT images remained almost blur-free, regardless of the speed, and they were similar to

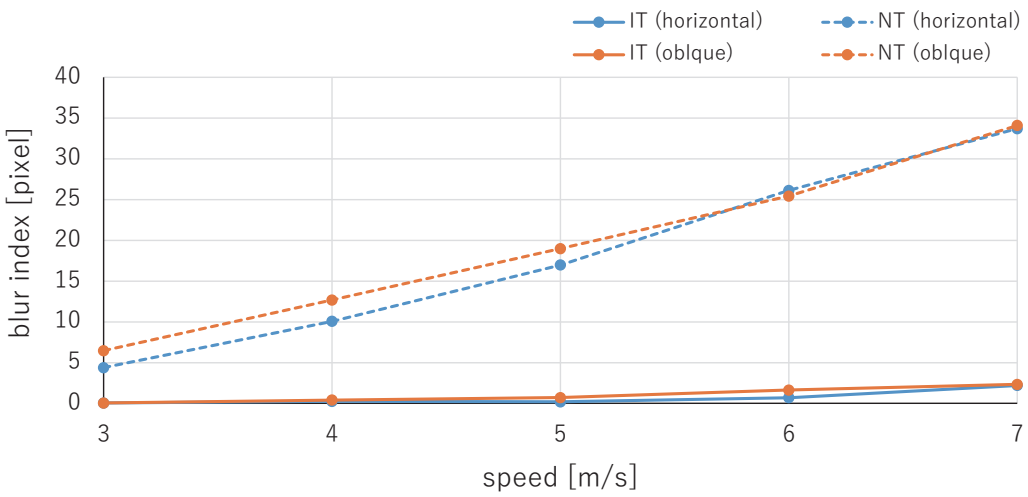


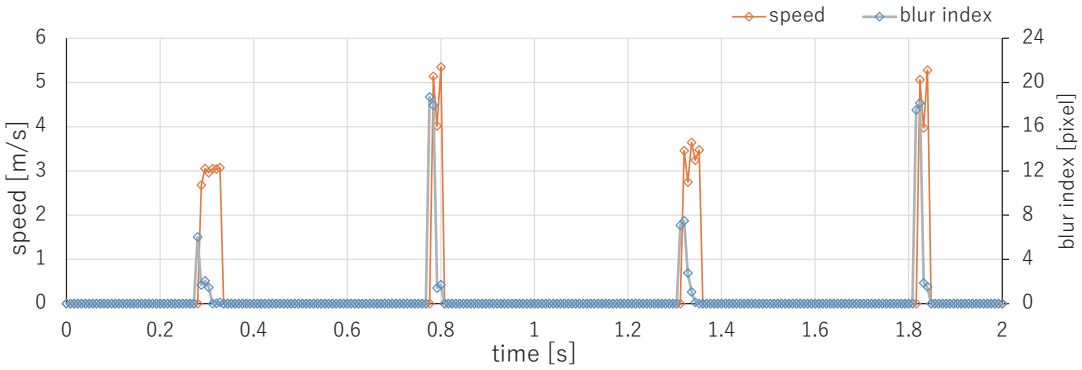
Figure 4.15: Motion blur indexes for table tennis balls thrown at different speeds.

the input images captured when the ball speed was 0 m/s, whereas the motion blur of the table tennis balls in the NT images increased in both their moving directions as their speed increased.

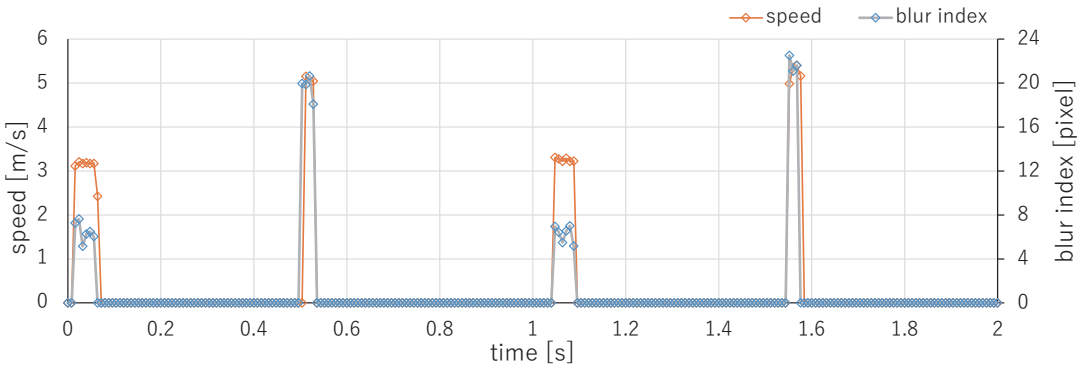
Figure 4.15 shows the relationship between the speed of a table tennis ball and its blur index $\Delta\lambda'$ for the IT and NT images. Considering the offset $\Delta\lambda_0 = 1.22$ pixel in the case of no motion, $\Delta\lambda'$ of those for 50 selected images, which were binarized with a threshold of 55, was averaged, in a manner similar to that in the experiment using circle-dot motion. As compared with the blur index $\Delta\lambda'$ for the NT images, the blur index $\Delta\lambda'$ for the IT images was remarkably low at all the speeds in the horizontal and oblique directions. The blur index $\Delta\lambda'$ for the IT images at 3, 4, 5, 6, and 7 m/s in the horizontal direction was 0.05, 0.26, 0.20, 0.70, and 2.20 pixel, respectively, which corresponds to 1.1, 2.6, 1.1, 2.7, and 6.5% of the respective value of $\Delta\lambda'$ for the NT images. The blur index $\Delta\lambda'$ for the IT images at 3, 4, 5, 6, and 7 m/s in the oblique direction was 0.05, 0.42, 0.71, 1.63, and 2.33 pixel, respectively, which corresponds to 0.8, 3.2, 3.8, 6.4, and 6.8% of the respective value of $\Delta\lambda'$ for the NT images. The blur index $\Delta\lambda'$ for the IT images showed a tendency to increase slightly when shooting a video of a table tennis ball thrown at 6 and 7 m/s in the horizontal and oblique directions. This is mainly because the speed of the table tennis ball was so much higher than the maximum motion-blur-free speed (5.21 m/s in the horizontal direction, 3.86 m/s in the vertical direction) that the moving

distance during an exposure time of 4 ms exceeded the upper limit of the movable range of the 2-DOF mirror-drive active vision system.

4.2.4 Table Tennis Ball Motion at Variable Speeds



(a) with tracking (IT)



(b) without tracking (NT)

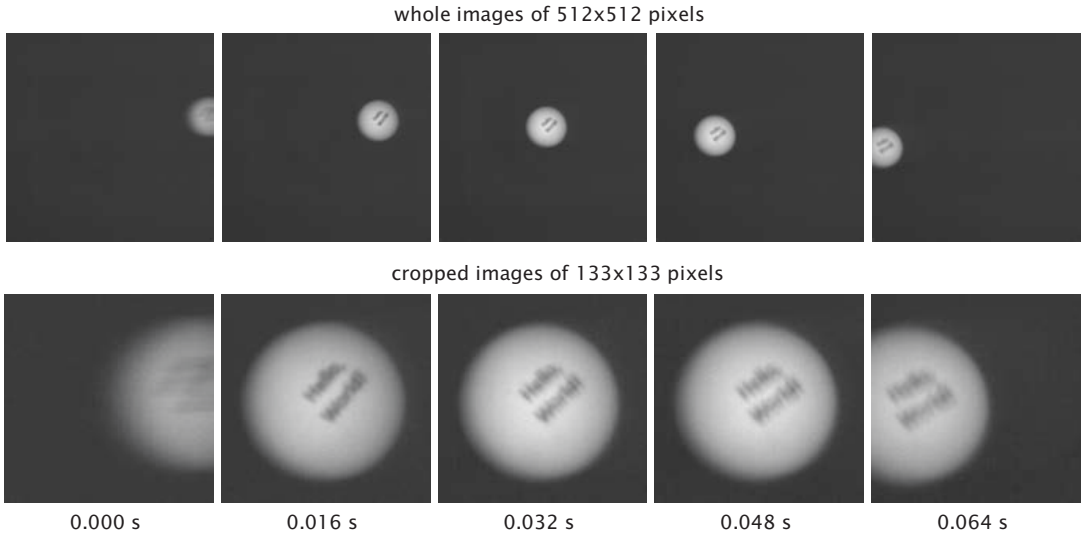
Figure 4.16: Estimated speed and motion-blur index when table tennis balls were thrown at variable speeds.

Next, we show the experimental results for the video of a table tennis ball identical to that used in the previous subsection, when the table tennis ball was alternately launched from the table tennis ball machine at different speeds of 3 m/s and 5 m/s at intervals of 0.5 s. Figure 4.16(a) shows the two-second temporal changes of the estimated speed and blur index $\Delta\lambda'$ for the IT images with frame-by-frame intermittent tracking, as compared with (b) those for the NT images when the table tennis balls were passing in front of the mirror-drive 2-DOF active vision system in a manner similar to that when capturing the IT images. Corresponding to the launching interval of a table tennis ball and its passing

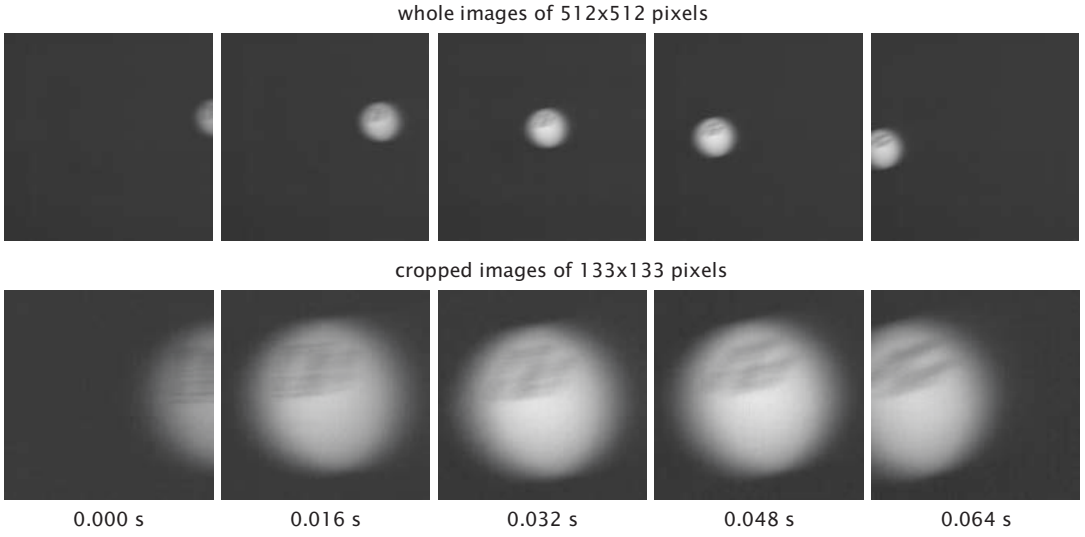
time duration over a whole image region of 512×512 pixels, the speeds of the table tennis balls in images were discontinuously estimated in time; the passing time durations were 66.7 and 40.0 ms when a table tennis ball was thrown at 3 m/s and 5 m/s, respectively; they correspond to the duration times for capturing eight and five frame images at 125 fps, respectively.

It can be seen that the ball speed was estimated as a pulse wave, in which a 3-m/s-amplitude pulse of 66.7-ms-width and a 5-ms-amplitude of 40-ms-width appear alternately at intervals of 0.5 s, and the blur index $\Delta\lambda'$ for the NT images also alternated between 7 and 20 pixels. The blur index $\Delta\lambda'$ for the IT images became a certain large value of 7 and 20 pixels exactly when the table tennis ball thrown at 3 m/s and 5 m/s appeared in the image, whereas it was remarkably reduced around 1 pixel dozens of milliseconds after its appearance in the image; this corresponds to the duration time for capturing two frame images at 125 fps. The latency in motion blur reduction is caused mainly by (1) the time delay in frame-by-frame intermittent tracking, involving a one-frame-delay in estimating the ball speed using image features computed at the previous frame and a one-frame delay in reflecting it to the pan-tilt actuation of the 2-DOF mirror-drive active vision system, and (2) underestimated speed exactly when the table tennis ball appears in the field of camera view because of its partial appearance at the right side of the images.

Figure 4.17 shows (a) a sequence of the images with frame-by-frame intermittent tracking, and (b) a sequence of the images without tracking when a table tennis ball with printed patterns, as illustrated in Figure 4.12(b), thrown at 3 m/s in the horizontal direction, was passing over the whole image region from right to left, taken at intervals of 16 ms; the upper images are the 512×512 input images, and the lower ones are the 132×132 images cropped from them so that the table tennis ball is located at their centers. It can be seen that the NT images are too heavily blurred to allow recognition of the letter patterns printed on the table tennis ball in all the frames. For the IT images, the input image was largely blurred at the start frame when the table tennis ball appeared at the right side in the image, whereas the blurring of the input images in all the remaining frames was reduced to the extent that the letter pattern of ”hello, world!” at the center of



(a) with tracking (IT)



(b) without tracking (NT)

Figure 4.17: Series of images captured when a table tennis ball was thrown in the horizontal direction.

the table-tennis ball can always be recognized.

Nevertheless, a two-frame delay remains in frame-by-frame intermittent tracking for motion blur reduction. Our system can capture less-blurred input images with a dozens-of-millisecond delay for a table tennis ball thrown at 8.0 m/s or less; its passing time over a whole image region of 512×512 pixels was larger than 24 ms for capturing three frame images at 125 fps.

4.3 Conclusion in this Chapter

In this chapter, we developed a motion-blur-free video shooting system based on a concept of the camera-driven frame-by-frame intermittent tracking, in which the control of the camera shutter state is alternated at a rate of hundreds of fps. The target's speed in images is controlled at zero during exposure; otherwise, the camera's position returns to its home position. Our system can capture 512×512 images of fast-moving objects at 125 fps with an exposure time of 4 ms without motion blur being incurred by controlling the pan and tilt directions of a mirror-drive 2-DOF active vision system using high-speed video processing. The system's performance was verified by conducting several experiments using fast-moving objects. We focused on motion blur reduction for moving objects in uniform backgrounds in this study, whereas both moving objects and static backgrounds can be clearly observed without blurring when the mirror speed is alternatively switched from the target object's speed to zero in frame-by-frame intermittent tracking so that image capturing "with tracking" (IT) and "without tracking" (NT) can be simultaneously conducted. Currently, the limited responses of the piezo actuators become the major bottleneck in frame-by-frame intermittent tracking at a higher frame rate; the duration time for back-to-home-control is 2 ms or more on our system, whereas the duration time for vision-based tracking control was set to approximately 4 ms, as shown in the angular displacements in Figure 4.4.

On the basis of these results, we plan to improve our motion-blur-free video shooting system by adapting it for video shooting of fast-moving objects in complex scenes with improved accuracy using fast general-purpose motion detection algorithms and faster frame-by-frame intermittent tracking using a free-vibration-type actuator such as a resonant mirror vibrating at hundreds or thousands of hertz, to apply our motion-blur-free video shooting system to highly magnified observations of fast-moving scenes in various applications, such as the precise inspection of products moving fast on a conveyor line and tunnel and road inspection from a vehicle moving at a high speed.

Chapter 5

Motion-Blur-Free High-Frame-Rate Video Camera using a Resonant Mirror

In this chapter, in order to verify the effectiveness of actuator-driven frame-by-frame intermittent tracking, which was extended from the camera-driven approach, we developed a test-bed system for motion-blur-free HFR video shooting using a resonant mirror.

5.1 Exposure Times and Vibration Amplitude

In motion-blur-free video shooting with actuator-driven frame-by-frame intermittent tracking, the nonlinear sinusoid trajectory with resonant vibration of a free-vibration-type actuator, which is segmented in the time range when the camera shutter is open, deviates from its approximate straight line more extensively as the camera exposure time increases. For motion-blur-free video shooting without lowering the incident light, it is important to determine a larger camera exposure time with consideration of the permissible deviation error in straight-line approximation, which corresponds to the degree of motion blur. In this section, we discuss how to determine parameters for camera exposure times and vibration amplitudes in actuator-driven frame-by-frame intermittent tracking on the basis of the numerical relationship between the segmented sinusoid trajectory and its approximate straight line.

As illustrated in Figure 5.1, the input image is captured at frame k with an exposure time τ by opening and closing the camera shutter at times $t_k^O = t_k - \tau/2$ and $t_k^C = t_k + \tau/2$, respectively. In this study, we assume that the center time of the camera exposure is set

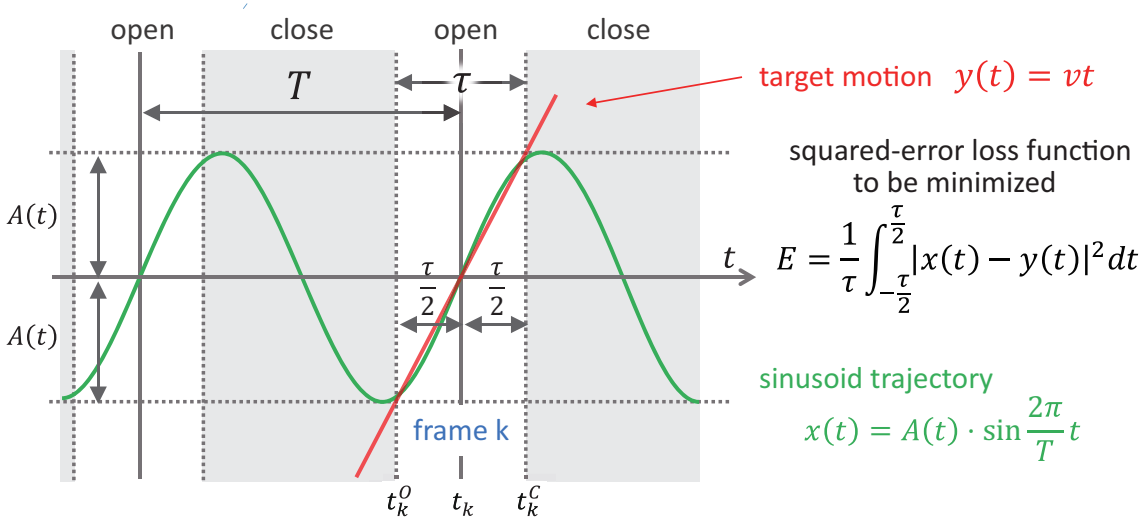


Figure 5.1: Sinusoid trajectory and its approximate straight line.

to $t_k = 2n\pi$ (n: integer) to synchronize with the sinusoid trajectory $x(t) = A\sin(2\pi/T)t$ so that the slope of a tangent to the sinusoid trajectory is maximum at time t_k. To track a target object moving at a speed of v in images when the camera shutter is open, we assume that the amplitude A of the sinusoid trajectory is so controlled that the straight line $y(t) = vt$ approximates the segmented sinusoid trajectory in the range of time t_k^O to time t_k^C. Here, we assume that the open and close times for camera exposure are $t_k^O = -\tau/2$ and $t_k^C = \tau/2$, respectively, by setting the center time to $t_k = 0$ for simplification, and the y-intercept of the approximate line is zero because the segmented sinusoid trajectory in the range of time t_k^O to time t_k^C is symmetric about the center time t_k. To estimate the amplitude A, we consider a minimization problem for the following squared-error loss function that can evaluate the deviation of the segmented sinusoid trajectory from the straight line where the target object moves,

$$
\begin{aligned}
E(A) &= \frac{1}{t_k^C - t_k^O}\int_{t_k^O}^{t_k^C} |x(t) - y(t)|^2\, dt = \int_{t_k^O}^{t_k^C} \left|A\sin\frac{2\pi}{T}t - vt\right|^2 dt, \\
&= \frac{1}{t_k^C - t_k^O}\left(A^2\int_{t_k^O}^{t_k^C} \sin^2\frac{2\pi}{T}t dt - 2Av\int_{t_k^O}^{t_k^C} t\sin\frac{2\pi}{T}t dt + v^2\int_{t_k^O}^{t_k^C} t^2 dt\right), \qquad (5.1) \\
&= A^2\cdot\left(\frac{1}{2} - \frac{T}{4\pi\tau}\sin\frac{2\pi\tau}{T}\right) - 2Av\cdot\left(-\frac{T}{2\pi}\cos\frac{\pi\tau}{T} + \frac{T^2}{2\pi^2\tau}\sin\frac{\pi\tau}{T}\right) + v^2\cdot\frac{\tau^2}{12}.
\end{aligned}
$$

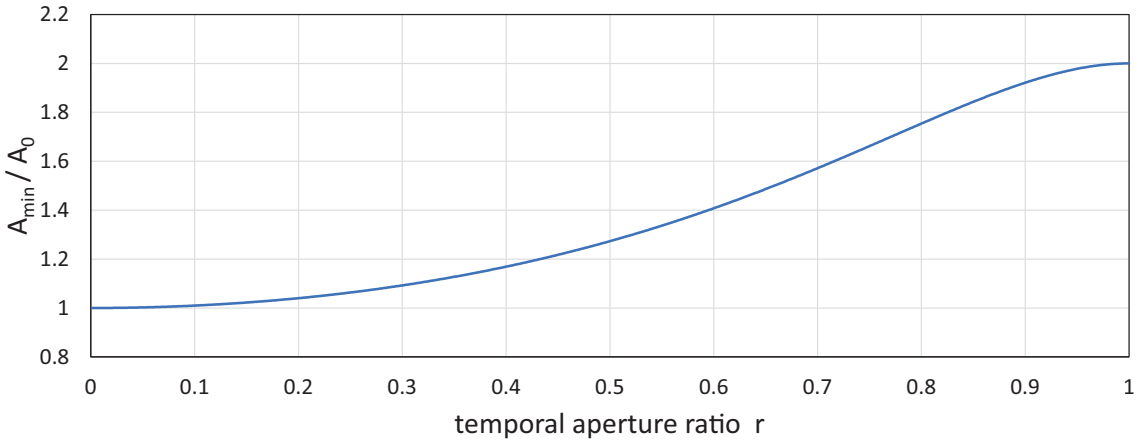


Figure 5.2: Relationship between temporal aperture ratio r and the amplitude ratio of A_{min} to A_0.

Solving the following equation such that the partial derivative of $E(A)$ with respect to A is zero,

$$\frac{\partial E}{\partial A} = 2A \cdot \left(\frac{1}{2} - \frac{T}{4\pi\tau}\sin\frac{2\pi\tau}{T}\right) - 2v \cdot \left(-\frac{T}{2\pi}\cos\frac{\pi\tau}{T} + \frac{T^2}{2\pi^2\tau}\sin\frac{\pi\tau}{T}\right) = 0, \tag{5.2}$$

the amplitude A_{min} can be derived as follows:

$$A_{min} = \frac{T^2\left(\sin\frac{\pi\tau}{T} - \frac{\pi\tau}{T}\cos\frac{\pi\tau}{T}\right)}{\pi^2\tau\left(1 - \frac{T}{2\pi\tau}\sin\frac{2\pi\tau}{T}\right)} \cdot v = \frac{T(\sin\pi r - \pi r\cos\pi r)}{\pi^2 r\left(1 - \frac{\sin 2\pi r}{2\pi r}\right)} \cdot v, \tag{5.3}$$

where $r = \tau/T$ is the temporal aperture ratio that indicates the ratio of the exposure time τ to the cycle time T of the sinusoid trajectory. Figure 5.2 shows the relationship between the temporal aperture ratio r and the amplitude ratio of A_{min} to A_0; A_0 is the slope of the tangent line of the sinusoid trajectory at time t_k when the exposure time τ approaches zero as follows:

$$A_0 = \lim_{\tau\to 0} A_{\min} = \frac{T}{2\pi}v. \tag{5.4}$$

Thus, the minimum value E_{min} of the squared-error loss function is obtained as follows:

$$E_{min} = \min_{A} E(A) = E(A_{min}) = \left(\frac{\tau^2}{12} - \frac{T^3}{\pi^3 \tau} \frac{(\sin \frac{\pi\tau}{T} - \frac{\pi\tau}{T} \cos \frac{\pi\tau}{T})^2}{\frac{2\pi\tau}{T} - \sin \frac{2\pi\tau}{T}} \right) v^2, \tag{5.5}$$

$$= \frac{T^2}{r} \left(\frac{r^3}{12} - \frac{1}{\pi^3} \frac{(\sin \pi r - \pi r \cos \pi r)^2}{2\pi r - \sin 2\pi r} \right) v^2. \tag{5.6}$$

Without actuator-driven frame-by-frame intermittent tracking, the squared-error loss E_{NT} in the range of time t_k^O to time t_k^C when observing a target object moving at speed v can be described as the value of the loss function when the amplitude of the sinusoid trajectory is $A = 0$, corresponding to no camera motion for tracking, as follows:

$$E_{NT} = E(A = 0) = \frac{1}{t_k^C - t_k^O} \int_{t_k^O}^{t_k^C} |vt|^2 \, dt = \frac{\tau^2}{12} \cdot v^2 = \frac{T^2 r^2}{12} \cdot v^2. \tag{5.7}$$

Considering the roots of the squared-error losses of E_{min} and E_{NT}, the relative error ratio ε is defined as follows:

$$\varepsilon = \frac{\sqrt{E_{min}}}{\sqrt{E_{NT}}} = \sqrt{1 - \frac{12}{\pi^3 r^3} \frac{(\sin \pi r - \pi r \cos \pi r)^2}{2\pi r - \sin 2\pi r}}, \tag{5.8}$$

where ε indicates the degree of motion blur reduction in video shooting with frame-by-frame intermittent tracking, compared with the deviation error in video shooting without tracking; $\varepsilon = f(r)$ is a monotonically increasing function of the temporal aperture ratio r, and motion blur is largely canceled when ε approaches zero. Thus, the temporal aperture ratio r can be expressed as a monotonically increasing function of the relative error ratio ε, which is independent of the cycle time T of the sinusoid trajectory and the target speed v, as follows:

$$r = f^{-1}(\varepsilon). \tag{5.9}$$

Figure 5.3 shows the relationship between the temporal aperture ratio r and the relative error ratio ε. Using the relationship between r and ε in Figure 5.3 as a look-up table,

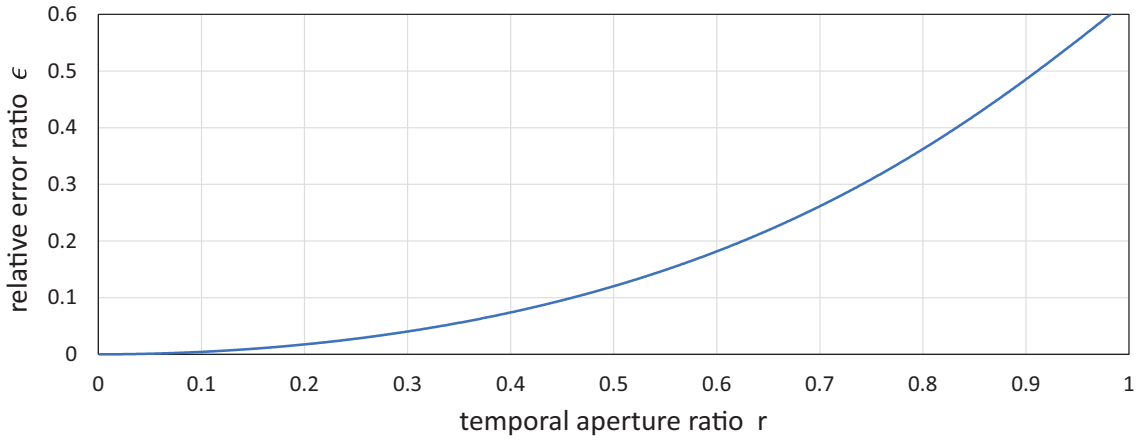


Figure 5.3: Relationship between temporal aperture ratio r and relative error ratio ε.

the camera shutter timings can be automatically determined in actuator-driven frame-by-frame intermittent tracking when the permissible degree of motion blur is initially given. For example, the relative error ratio ε is permissible up to 1%, 5% and 10%, respectively, and the upper-limit values of the allowable temporal opening ratios are $r(0.01) = 0.151$, $r(0.05) = 0.333$, and $r(0.1) = 0.460$, respectively. Especially when the exposure time is constant, the open and close times for camera exposure can be determined independently from the time-varying amplitude A of the sinusoid trajectory, which is controlled so as to cancel the apparent speed of the target object when the camera exposure is open, and these signals are generated in synchronization with the external synchronization signal from a free-vibration-type actuator.

5.2 Developed Motion-blur-free HFR Video Camera System using a Resonant Mirror

Figure 5.4 shows the overview of the test-bed system. The test-bed system consists of (1) a motion-blur-free HFR camera system with a resonant mirror and (2) a high-speed belt-conveyor system that can convey target objects to be observed at various speeds.

The motion-blur-free HFR camera system consists of a high-speed vision platform develop based on FASTCAM SA-X2 (Photron, Tokyo, Japan) with an F-mount

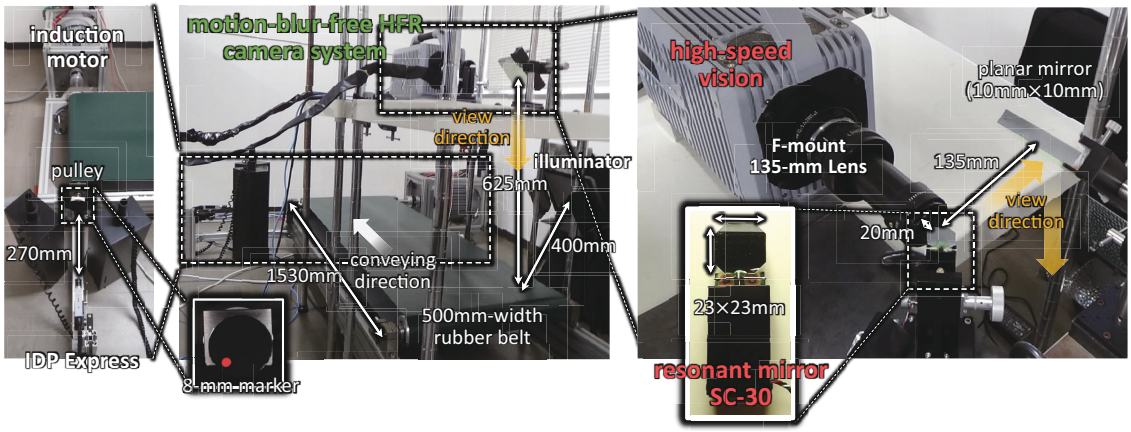


Figure 5.4: Overview of the test-bed system for motion-blur-free video shooting.

135-mm lens (Rodagon 135 mm F5.6, Qioptiq Photonics, Hamble-le-Rice, UK), a resonant mirror (SC-30, Electro-Optical Products, Ridgewood, NY, USA), a function generator (AFG1022, Tektronix, Beaverton, OR, USA) and a personal computer (PC) with an ASUSTeK P6T7 WS Supercomputer main board, Intel Core i7 960 3.20-GHz CPU, 6-GB memory, Windows 7 Professional 32-bit OS and a D/A board (PEX-340416, Interface, Hiroshima, Japan). FASTCAM SA-X2 has a 12-bit 1024 × 1024 CMOS image sensor, the sensor size and pixel size of which are 20.48 × 20.48 mm and 20 × 20 μm, respectively, and it can capture and record 1024 × 1024 images at 12,500 fps in inner memories with a 256-parallel analog output at a 65-MHz clock. The camera parameters such as exposure time can be set through a Gb Ethernet from external systems, whereas the shutter timings to open the camera exposure can be directly determined by an external trigger signal. SC-30 is a resonant mirror system that can vibrate a 23 × 23-mm-size glass mirror in the pan direction at its resonant frequency of 750 Hz. The amplitude of its sinusoid trajectory can be controlled in the range of 0.0025 to 0.5 degrees by providing a voltage command in the range of 0 to 5 V to its automatic gain control driver; the voltage command was outputted from the D/A board mounted on the PC. A transistor-transistor-logic (TTL) signal is externally outputted at the cycle time of its resonant vibration, and it was used to determine the shutter timings to open the camera exposure at 750 Hz. The function generator AFG1022 was used to adjust the delay time in the external TTL signal so as to synchronize the center time of the open exposure with the vibration center

of the sinusoid trajectory in this study. The PC was mainly used to control the vibration amplitude of the resonant mirror system for motion-blur-free video shooting.

The high-speed belt-conveyor system was installed 625 mm below a planar mirror of 100 × 100 mm size; the planar mirror was located 135 mm on the right side of the resonant mirror to change the direction of the camera view to the vertical direction for target objects horizontally-moving on the belt-conveyor system; they were observed under the lighting with an LED illuminator (VLP-10500XP, LPL, Saitama, Japan), which was installed 400 mm diagonally upward from the belt-conveyor system. On the belt-conveyor system, target objects attached to a 500-mm-width rubber belt can move forward with rotations of 80 mm-diameter pulleys, one of which was the drive pulley powered with a three-phase induction motor (SF-PRV-3.7KW-4P-200V, Mitsubishi Electric, Japan). The length between pulleys was set to 1530 mm. The induction motor was controlled by an inverter (WJ200-037LF, Hitachi Industrial Equipment Systems, Tokyo, Japan), and the conveying speed of the belt-conveyor system can be set in the range of 0 to 7.55 m/s by providing a voltage command in the range of 0 to 10 V to the inverter. The rotation speed of the induction motor was measured by a high-speed vision system IDP Express [74]; the rotation speed was computed by extracting the position of an 8 mm-diameter marker attached on a rotational axis with real-time video processing of 512 × 512 images at 2000 fps on IDP Express. Thus, the conveying speed of the belt-conveyor system can be simultaneously estimated for motion-blur-free video shooting at 2000 Hz on the PC, on which IDP Express was mounted for controlling the vibration amplitude of the resonant mirror.

With this test-bed system, 1024 × 1024 input images were captured at 750 fps, with a frame interval of 1.333 ms and the exposure time of 0.33 ms, respectively, in synchronization with the external trigger signal from the resonant mirror, which was dependent on its resonant frequency. The intensity of incident light on the image sensor was not different from that while shooting a video without a resonant mirror. We confirmed that only 2.2% of the intensity of incident light was lost owing to the 23 × 23 mm mirror in the presented setup, compared with that in video shooting without a resonant mirror. The temporal aperture ratio in frame-by-frame intermittent tracking was $r = 0.25$, and the

relative error ratio $\varepsilon = 0.028$; this corresponded to the relationship between r and ε in Figure 5.3. A target object on the belt plane was apparently distant $R = 760$ mm from the center of the resonant mirror. Considering the twice of the mirror angle of the resonant mirror, the sinusoid trajectory of the mirror angle $\theta(t) = A_\theta \sin 2\pi t/T$ was projected on the belt plane as:

$$x(t) = 2A_\theta R \sin \frac{2\pi}{T} t, \tag{5.10}$$

where it is assumed that $\theta(t)$ is small. From Equation (5.3), the vibration amplitude A_θ of the resonant mirror to minimize the squared-error loss function described in Equation (5.2) can be expressed as a function of the speed v of a target object by substituting $T = 1.333$ ms, $r = 0.25$ and $R = 760$ mm as follows:

$$A_\theta = \frac{1}{2R} \frac{T(\sin \pi r - \pi r \cos \pi r)}{\pi^2 r \left(1 - \dfrac{\sin 2\pi r}{2\pi r}\right)} \times v = c_{min} \times v, \tag{5.11}$$

where $c_{min} = 1.485 \times 10^{-4}$ (rad·s/m) $= 8.506 \times 10^{-3}$ (deg·s/m). The vibration amplitude of the resonant mirror was controlled with sensor feedback so that the apparent scan speed with the resonant mirror on target objects on the belt was always matched with the conveying speed measured by the high-speed vision system IDP Express. An image region of 1024×1024 pixels corresponded to the 104×104 mm area on the belt of the belt-conveyor system, and 0.10 mm corresponded to one pixel. The maximum permissible speed for target objects to be observed that can guarantee the efficiency of frame-by-frame intermittent tracking was determined theoretically by the ratio of 0.5 deg to 0.33 ms, which were the maximum movable angle of the resonant mirror and the duration time of the exposure time, respectively; the maximum angular speed was 2.34×10^3 deg/s considering that the variation of the view angle via the mirror corresponds to twice that of the mirror angle. Thus, the displacement of 95.6 pixels in the x direction on the image sensor was permissible during the open exposure of 0.33 ms, and the maximum permissible apparent speed on the image sensor was 2.90×10^5 pixel/s. This value

corresponded to the maximum permissible speed of 29.4 m/s for objects to be observed on the belt of the belt-conveyor system, whereas the maximum conveying speed of the belt-conveyor system was 7.55 m/s.

5.3 Preliminary Experiments

5.3.1 Relationship between Drive Voltage and Vibration Amplitude

Firstly, we conducted a preliminary experiment to verify the relationship between the drive voltage to the resonant mirror and its angular displacement. To measure the angular displacement, a laser beam spot was redirected by the resonant mirror, and the locations of the beam spots projected on a screen at a distance of 1375 mm from the resonant mirror were extracted offline by capturing an HFR video of 384 × 56 pixels at 100,000 fps with the exposure time of 8.98×10^{-3} ms. Figure 5.5 shows the angular displacement for 4 ms when the drive voltage to the resonant mirror was set to 0.0, 1.0, 2.0, 3.0, 4.0 and 5.0 V. The angular displacement was sinusoidally changed at a frequency of 750 Hz, and its amplitude increased in proportion with the drive voltage. Figure 5.6 shows the relationship between the drive voltage and the averaged vibration amplitude of the angular displacement for 1 s, corresponding to 750 cycle times of the 750-Hz vibration, when the drive voltage varied in the range of 0.0 to 5.0 V at steps of 0.2 V. The vibration amplitude linearly varied with the amplitude of the drive voltage, whereas there was a slight offset around 0 V; the relationship between the drive voltage V (V) and the vibration amplitude A (deg) can be linearly approximated as $A = 0.0368V + 0.0026$. Figure 5.7 shows the relationship between the drive voltage and the standard deviation of the vibration amplitude in the duration time of 1 s. In the figure, the relative ratio of the standard deviation to the averaged vibration amplitude was also plotted. When the drive voltage was 5.0 V, the averaged vibration amplitude and its standard deviation were 0.188 and 3.35×10^{-4} deg, respectively. The standard deviations had similar values around 3×10^{-4} deg at all the drive voltages, and the relative ratio decreased in proportion to the drive voltage; the relative ratio was 0.18% when the drive voltage was 5 V.

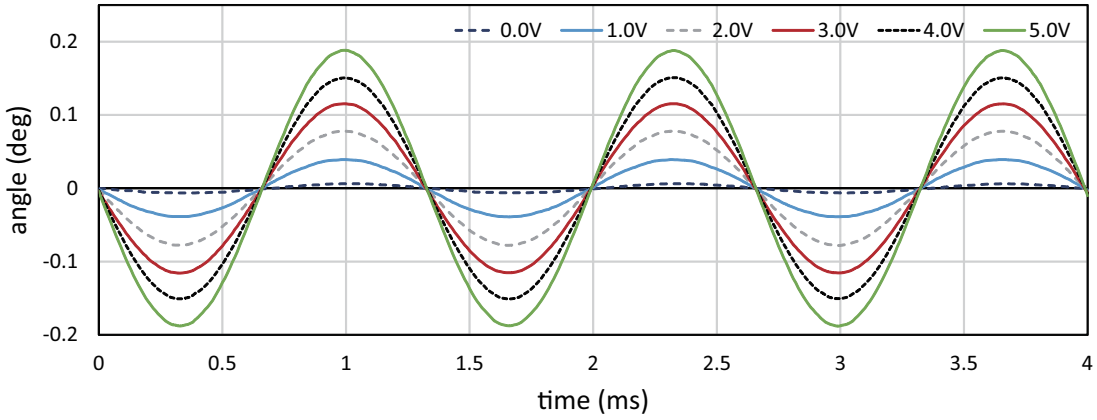


Figure 5.5: Relationship between drive voltage and angular displacement.

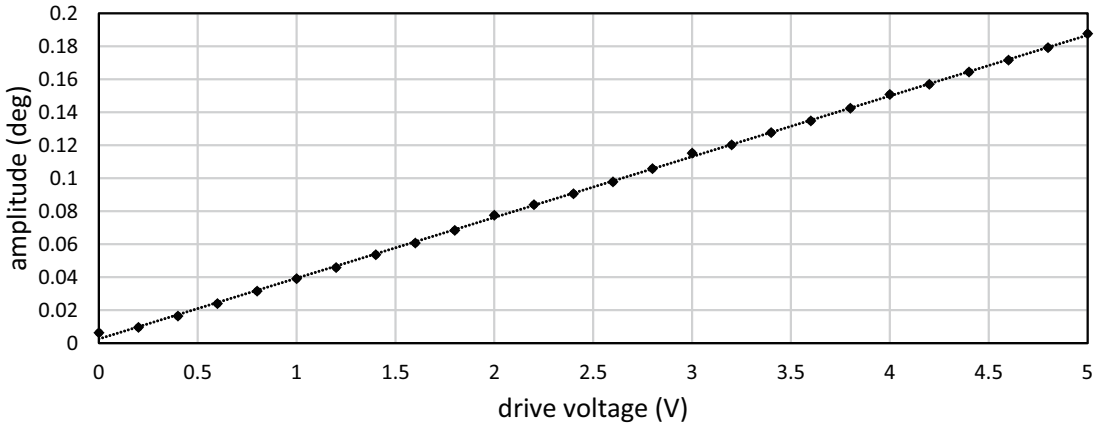


Figure 5.6: Relationship between drive voltage and vibration amplitude.

5.3.2 Step Responses of Vibration Amplitude

Next, we conducted an experiment to verify the response time of the vibration amplitude of the resonant mirror when a step drive voltage is commanded to the resonant mirror. In a similar environment as that in the previous subsection, the vibration amplitude of the resonant mirror is measured by capturing an HFR video for the laser beam spots projected on a screen; 384 × 265 images were captured for 5 s at 750 Hz with the exposure time of 0.05 ms in synchronization with the timing when the angular displacement of the resonant mirror was at the maximum. Figure 5.8 shows the step response of the vibration amplitude when the drive voltage of (a) 1 V and (b) 3 V is simultaneously switched to the different target voltage in the range of 0 to 5 V at time $t = 0$. Figure 5.9

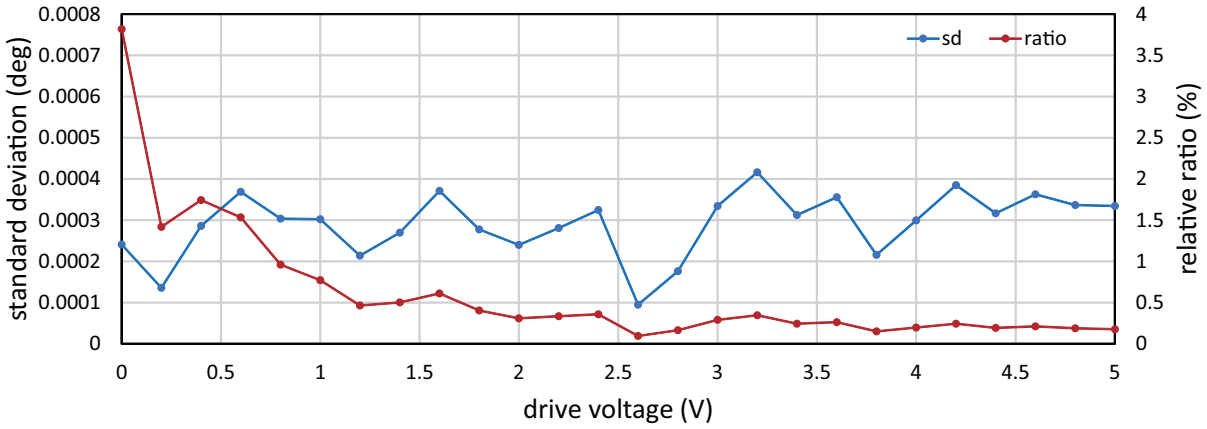
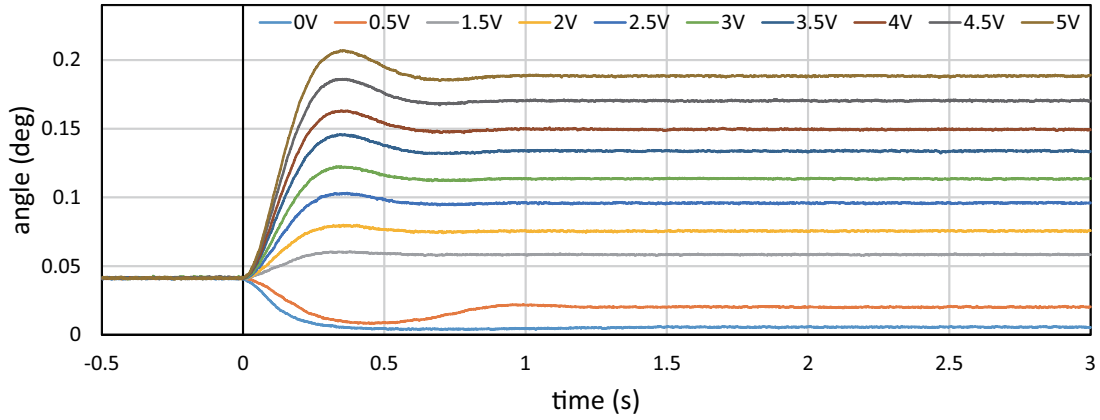
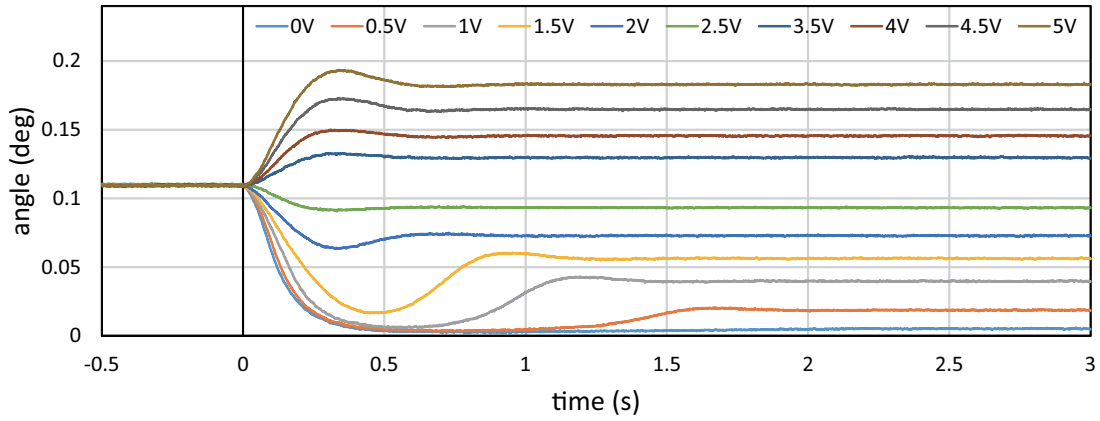


Figure 5.7: Relationship between drive voltage and the standard deviation of vibration amplitude.

shows the rise time (from 10 to 90%), the delay time (to 50%) and the settling time (within 5 %) of the vibration amplitude when analyzing the step responses in Figure 5.8. At all target voltages except 0 V in (a) 1 V and (b) 3 V, the rise times and the delay times had similar values of 0.16 s and 0.12 s, respectively. However, the settling times were much larger than these parameters. There was a distinct tendency of hysteresis that the settling time was around 0.50 s for all cases when the drive voltage increased, whereas it became larger when the drive voltage largely decreased. Comparing with the 750-Hz free vibration of the resonant mirror, the dynamic response of the vibration amplitude is so slow and hysterical that the vibration amplitude cannot be quickly controlled for motion blur reduction when the speed of a target object to be tracked is quickly time-varying, whereas vibration amplitude control functions well in motion-blur-free video shooting for a target object moving at a large, but slightly time-varying speed in many applications such as product inspection on a factory line and road inspection from a moving vehicle.



(a)



(b)

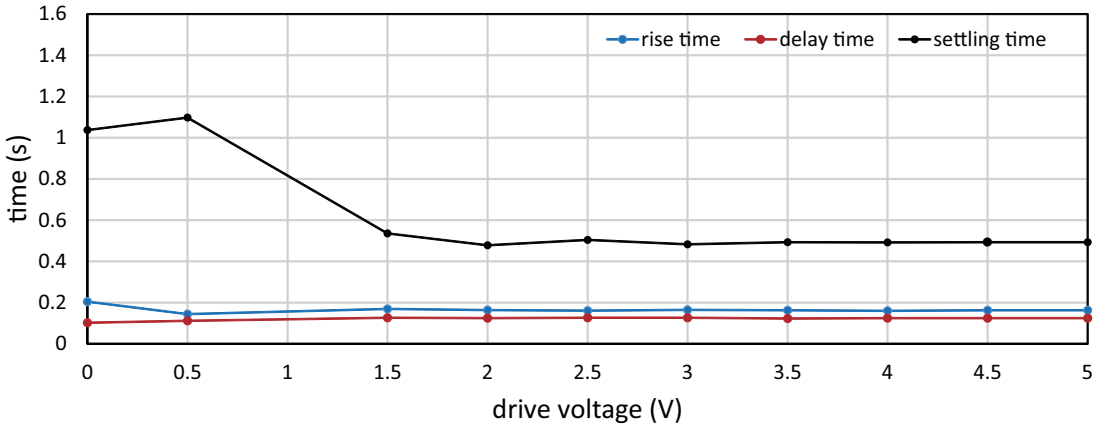
Figure 5.8: Step response of vibration amplitude: (a) 1 V; (b) 3 V.

5.4 Video Shooting Experiments

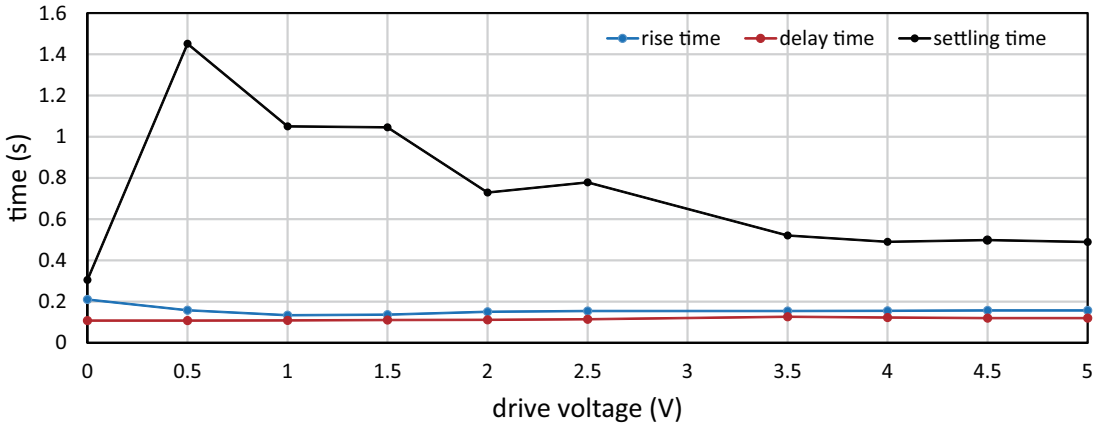
5.4.1 Video Shooting without Amplitude Control for Circle-Dots Moving at Constant Speeds

Next, we conducted video shooting experiments for a patterned object attached on the moving belt of the test-bed system to verify the relationship between the speed of a target object and its motion blur when the vibration amplitude of the resonant mirror was set to a constant value. Figure 5.10 shows the patterned object to be observed; a circle-dot pattern on which 4 mm-diameter circle-dots were black-printed at vertical and horizontal intervals of 11 mm and 7 mm, respectively.

Figure 5.11 shows the 215 × 215 images cropped from the 1024 × 1024 input im-



(**a**)



(**b**)

Figure 5.9: Dynamic response parameters of vibration amplitude: (a) 1 V; (b) 3 V.

ages of the circle-dot pattern moving with the motor command of 0.0, 1.0, 2.5, 4.0 and 6.5 m/s to the conveyor system when the vibration amplitude of the resonant mirror was set to 0.0000, 0.0063, 0.0202, 0.0391 and 0.0572 deg; 0.0000 deg corresponded to no vibration of the resonant mirror, and 0.0063, 0.0202, 0.0391 and 0.0572 deg corresponded to the voltage command of 0.0, 0.5, 1.0, 1.5 V to the control driver of the resonant mirror, respectively. The patterns moving with the motor command of 0.0, 1.0, 2.5, 4.5 and 6.5 m/s were captured without motion blur when the vibration amplitude of the resonant mirror was 0.0000, 0.0063, 0.0202, 0.0391 and 0.0572 deg, respectively. This tendency corresponded to that the squared-error loss functions when the vibration amplitude of the resonant mirror was 0.0000, 0.0085, 0.0213, 0.0383 and 0.0553 deg were minimized in

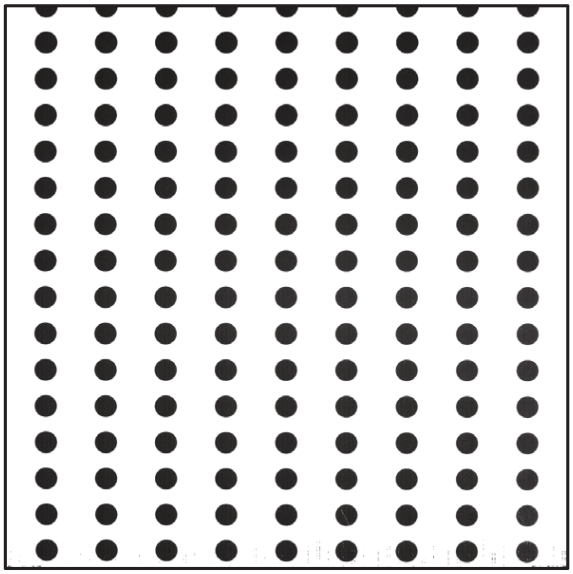


Figure 5.10: Circle-dot pattern to be observed at 1024×1024 pixels image.

observing a target object moving with the motor command of 0.0, 1.0, 2,5, 4.5 and 6.5 m/s, respectively, according to Equation (5.11). The image degradation with motion blur in the horizontal direction became larger as the object speed deviated from its desired speed for motion blur reduction, which was determined by the vibration amplitude of the resonant mirror.

For a circle-dot pattern, the blur index $\lambda_{dot} = \lambda_x - \lambda_x^0$ was introduced; λ_x represents the length of the x-axis of the approximated ellipse of the circle dot in the image, and λ_x^0 is the value of λ_x in observing a circle-dot at a fixed location in the case of no vibration of the resonant mirror. The index λ_{dot} increases as the motion blur in the horizontal direction becomes larger in the image, and it becomes zero when the circle-dot has no motion. λ_x was estimated offline by computing its zero-, first- and second-order moment features for the circle-dot region in the 424×424 image cropped from the input image; a single circle-dot was located at the center of the cropped image. The circle-dot region was extracted by binarization with a threshold of 2600.

Figure 5.12 shows the relationship between the speed of a circle-dot and its blur index λ_{dot} when the target objects moved with the motor command to the conveyor system in the range of 0.0 to 7.5 m/s at intervals of 0.5 m/s. In the figure, the blur indexes λ_{dot} were averaged by those for 25 selected dots in two images, and they were plotted when the vi-

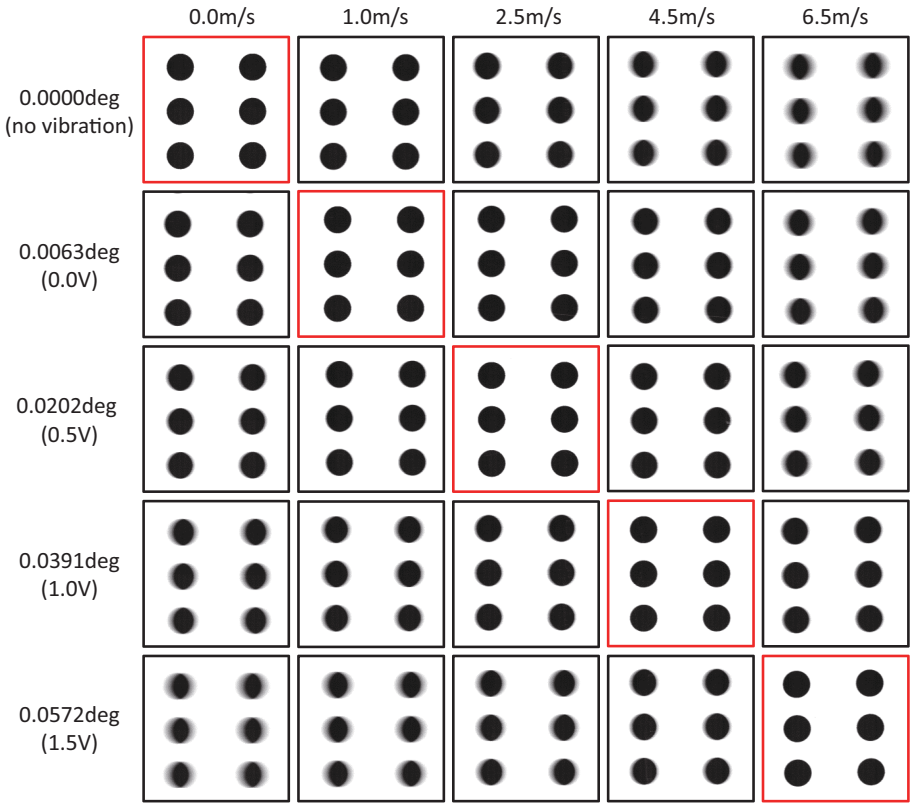


Figure 5.11: Images of a circle-dot pattern when video shooting without amplitude control.

bration amplitude of the resonant mirror was 0.0000, 0.0063, 0.0202, 0.0391, 0.0572 and 0.0776 deg. When the vibration amplitude was 0.0000, 0.0063, 0.0202, 0.0391 and 0.0572 deg, the blur index λ_{dot} had a minimum value when the motor command was 0.0, 1.0, 2.5, 4.5 and 6.5 m/s, respectively; these motor commands were around the desired speeds for motion blur reduction, 0.00, 0.74, 2.38, 4.60 and 6.72 m/s, which were determined by the vibration amplitude 0.0000, 0.0063, 0.0202, 0.0391 and 0.0572 deg, respectively. The desired speed for motion blur reduction when the vibration amplitude of the resonant mirror was 0.0776 deg, corresponding to the voltage command of 2.0 V to its control driver, was 9.12 m/s, and it was larger than the maximum conveying speed of 7.55 m/s on the belt-conveyor system. Thus, there were no local maximum/minimum values when the vibration amplitude was 0.0776 deg in Figure 5.12. These experimental results indicate that we can reduce motion blur in video shooting when the object speed corresponds to the desired speed for motion blur reduction, which is determined by the vibration amplitude

of the resonant mirror.

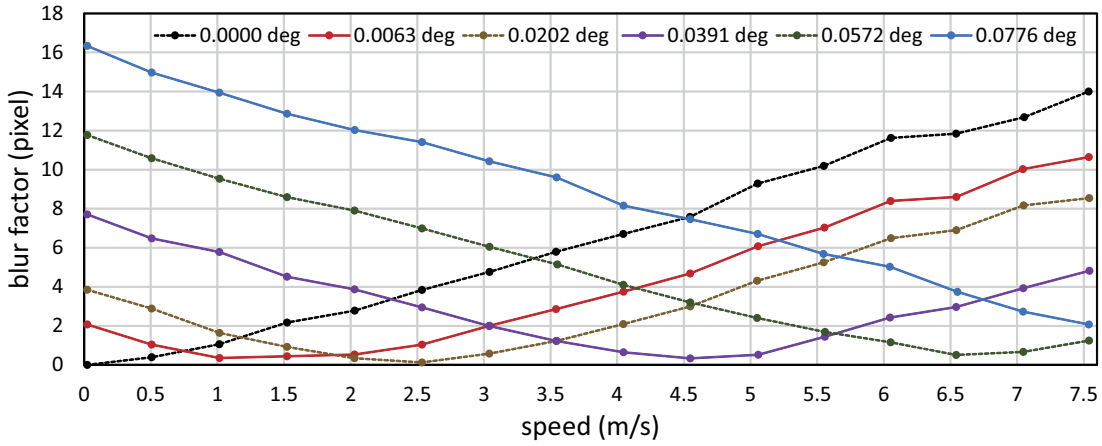


Figure 5.12: Blur indexes λ_{dot} when video shooting without actuator control.

5.4.2 Video Shooting with Amplitude Control for Circle-Dots Moving at Constant Speeds

Next, we conducted video shooting experiments for a fast-moving patterned object when the vibration amplitude of the resonant mirror was controlled in proportion to the object speed on the belt, which was estimated at 2000 fps by IDP Express, so that motion blurs in input images were reduced by frame-by-frame intermittent tracking. The circle-dot pattern identical to that used in Section 5.4.1 was observed in the experiments.

Figure 5.13 shows the 215 × 215 images cropped from the 1024 × 1024 input images of the circle-dot pattern when the patterned objects moved with the motor command to the conveyor system of 0.0, 1.5, 3.0, 4.5, 6.0 and 7.5 m/s. The input images captured when the vibration amplitude of the resonant mirror was controlled with sensor feedback (IT, with tracking) were compared with those captured when no vibration of the resonant mirror (NT, without tracking). As observed in Figure 5.13, the NT images became increasingly blurred in the horizontal direction as the object speed increased, whereas the IT images were non-blurred for all the speeds. Figure 5.14 shows the relationship between the speed of a circle-dot and its blur index λ_{dot} in video shooting the IT images with amplitude control and the NT images without amplitude control when

the target objects moved with the motor command to the conveyor system in the range of 0 to 7.5 m/s at intervals of 0.5 m/s. In the figure, the vibration amplitudes of the resonant mirror in video shooting the IT images with amplitude control were also plotted. The index λ_{dot} was computed offline in a similar manner as in the previous section. In Figure 5.14,the vibration amplitudes of the resonant mirror were controlled for motion blur reduction in proportion to the object speed in video shooting with amplitude control, whereas there remained a slight non-zero offset in the vibration amplitude when the motor command to the conveyor system was 0 m/s. In Figure 5.14, the blur index λ_{dot} for the IT images was remarkably low, comparing with that for the NT images. The blur index λ_{dot} for the IT images when the motor command was 0.0, 1.5, 3.0, 4.5, 6.0 and 7.5 m/s was 1.50. 0.07, 0.00, 0.02, 0.09 and 0.33 pixel, respectively, whereas that for the NT images was 0.00, 1.57, 4.45, 7.25, 10.36 and 13.42 pixel, respectively. In the experiments, the object speed was 7.55 m/s or less, which was smaller than the maximum permissible motion-blur-free speed of 29.4 m/s in the horizontal direction, and our actuator-driven frame-by-frame intermittent tracking method remarkably reduced motion blurs of the circle-dot pattern moving at high speed. When the motor command to the conveyor system was 1.0 m/s or less, the blur index λ_{dot} in video shooting with actuator control was slightly larger than that in video shooting the object moving with the motor command of 1.5 m/s or more. This is because the resonant mirror could not control its vibration amplitude around 0 deg due to friction hysteresis, and there still remained small vibration as the relationship between the drive voltage and vibration amplitude was described in Section 5.3.1.

5.4.3 Video Shooting with Amplitude Control for Patterned Objects at Variable Speeds

Next, we show the experimental results in video shooting with amplitude control for a checkered pattern, when the object speed varied in the range of 0 to 7.55 m/s. In the experiments, the motor command for the belt-conveyor system was set to the following trajectory: the motor command that determines the object speed started to increase from

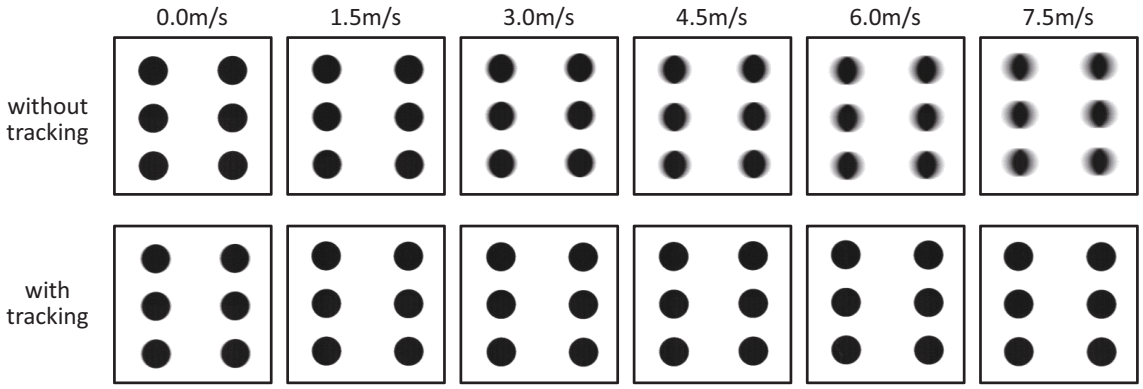


Figure 5.13: Images of a circle-dot pattern when video shooting with amplitude control.

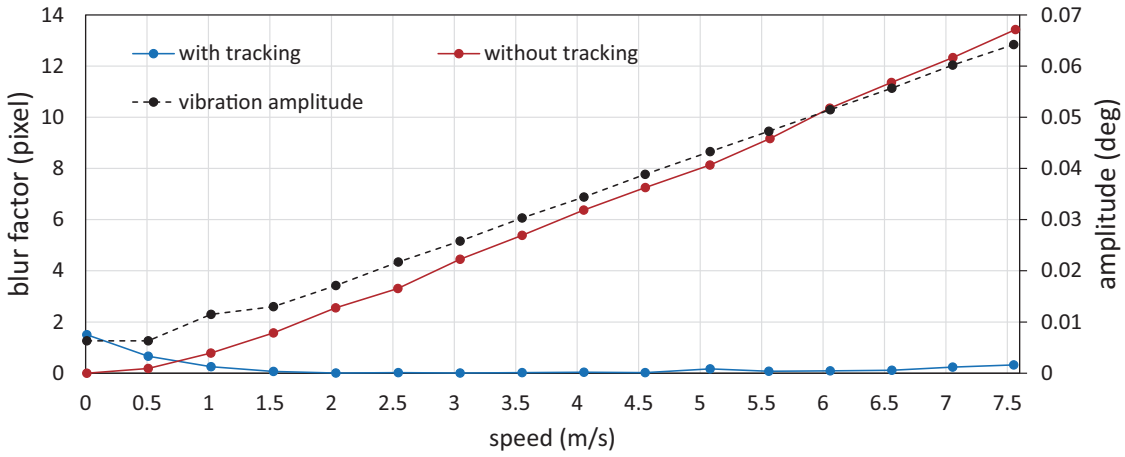


Figure 5.14: Blur indexes λ_{dot} for a circle-dot pattern and the vibration amplitude of resonant mirror when video shooting with actuator control.

0 m/s at time t = 4.3 s and reached the maximum conveying speed of 7.55 m/s at time t = 8.0 s. After keeping the maximum speed during 8.0 s for t = 8.0 to 16.2 s, it started to decrease and reached 0 m/s at time t = 26.0 s. Figure 5.15 shows a checkered pattern on which 2 × 2 mm squares of alternating black and white were printed.

Figure 5.16 shows (a) the measured object speed and the vibration amplitude of the resonant mirror and (b) the blur index λ_{edge} when the IT images for t = 0.0 to 30.0 s were captured with frame-by-frame intermittent tracking. The blur index $\lambda_{edge} = E_{av}/I_{ave}$ was introduced; I_{ave} and E_{ave} are the averaged values of the image intensities $I(x, y)$ and edge intensities $E(x, y)$, respectively, for the 824 × 824 image center-cropped from a 1024 × 1024 input image. The index λ_{edge} decreases as the motion blur becomes larger in the

image, because the edge intensities in the moving direction are degraded due to motion blur. The edge intensities were computed as follows:

$$E(x, y) = \sqrt{|I(x + 1, y) - I(x, y)|^2 + |I(x, y + 1) - I(x, y)|^2}. \tag{5.12}$$

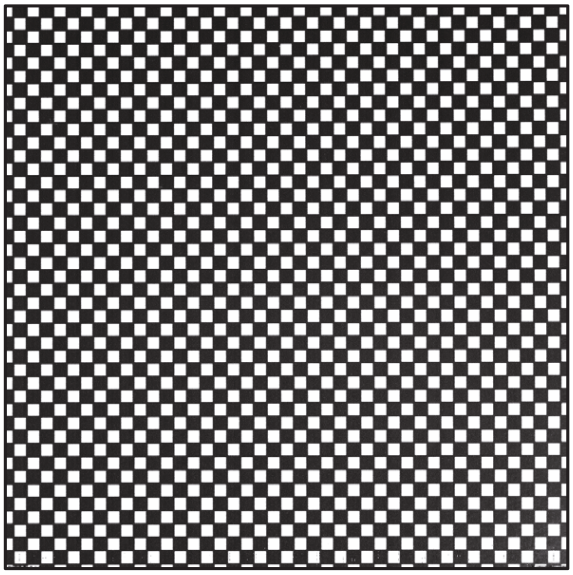
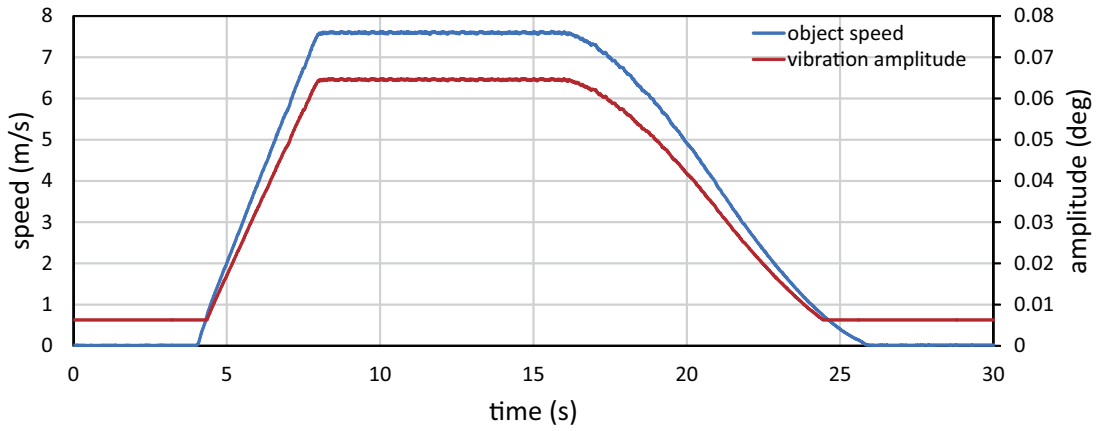
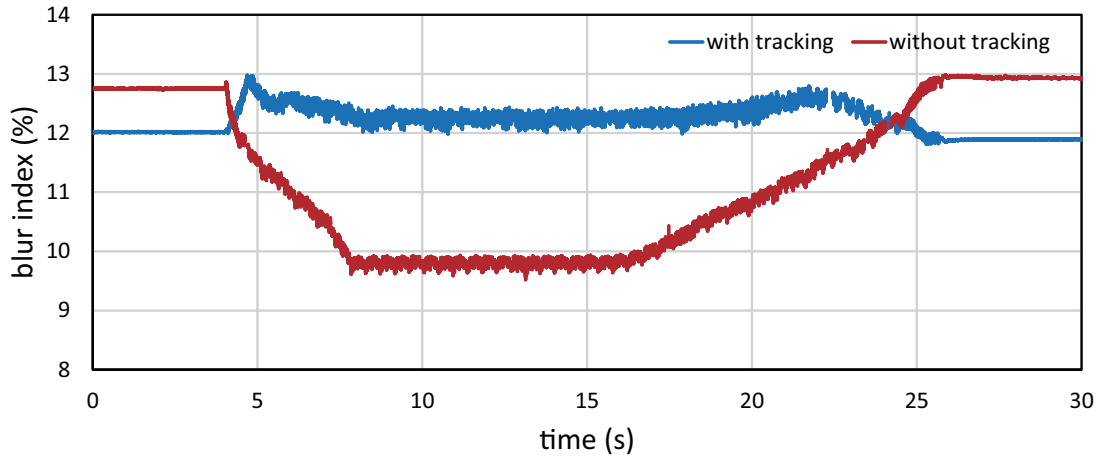


Figure 5.15: Checkered pattern to be observed at 1024×1024 pixels image.

For comparison, Figure 5.16b shows the blur index λ_{edge} when the NT images were captured with no vibration of the resonant mirror; the checkered pattern moved in a manner similar to the captured IT images. Figure 5.16b shows that the value of λ_{edge} for the IT images was almost constant in the range of 11.8 to 13.0% when the speed of the checkered pattern varied in time, whereas the value of λ_{edge} for the NT images varied considerably in the range of 9.5 to 13.0%, depending on the speed of the checkered pattern. The blur indexes λ_{edge} for the NT images were larger than those for the IT images when the measured object speed was 0 m/s due to the non-zero offset in the vibration amplitude of the resonant mirror as illustrated in Figure 5.16a. There were fluctuations in the blur indexes λ_{edge} for both the IT and NT images when the checkered pattern was moving because the average values of the edge intensities were slightly varied depending on the apparent location of the checkered pattern in the images. The dynamic response of the vibration amplitude of the resonant mirror was not so quick compared with its 750-



(**a**)

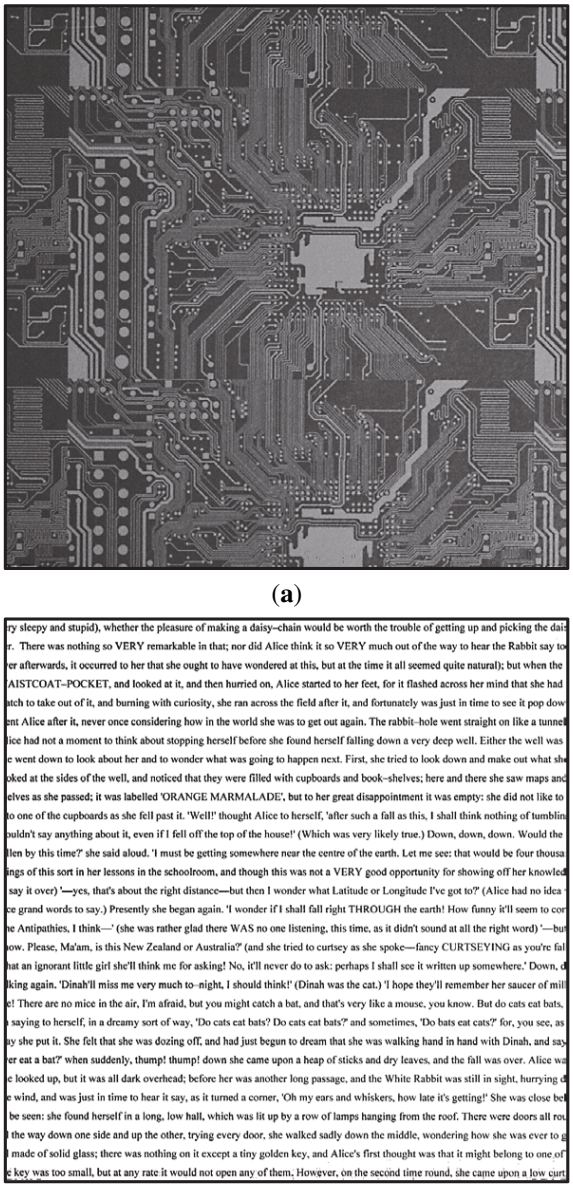


(**b**)

Figure 5.16: Experimental results for a checkered pattern moving at variable speeds: (a) object speed and vibration amplitude of resonant mirror; (b) blur indexes λ_{edge}.

Hz free vibration as described in Section 5.3.2, whereas the vibration amplitude control functioned well for motion blur reduction with frame-by-frame intermittent tracking in video shooting a target object moving at a large, but slightly time-varying speed on the belt-conveyor system, because the dynamic response of the conveyor's speed was slower than that in the vibration amplitude control of the resonant mirror.

To verify our actuator-driven frame-by-frame intermittent tracking with amplitude control for complex patterned objects, we experimented with (a) the printed pattern of an electronic board of 54 × 84 mm in size with 0.25 mm-width wiring patterns and (b)



(**a**)

(**b**)

Figure 5.17: Electronic board pattern and book page pattern to be observed at 1024×1024 pixels image: (a) electronic board pattern; (b) book page pattern.

the printed pattern of a book page with many 2-mm letters as illustrated in Figure 5.17; these patterns were attached on the belt of the belt-conveyor system and moved at variable speeds in a manner similar to the checkered pattern. Figures 5.18 and 5.19 show (a) the 323 × 323 IT images cropped from the 1024 × 1024 input images of the electronic board pattern and (b) the 323 × 323 IT images of the book page pattern, when the object speed was 0.0, 2.5, 5,0 and 7.5 m/s, compared with the NT images captured when

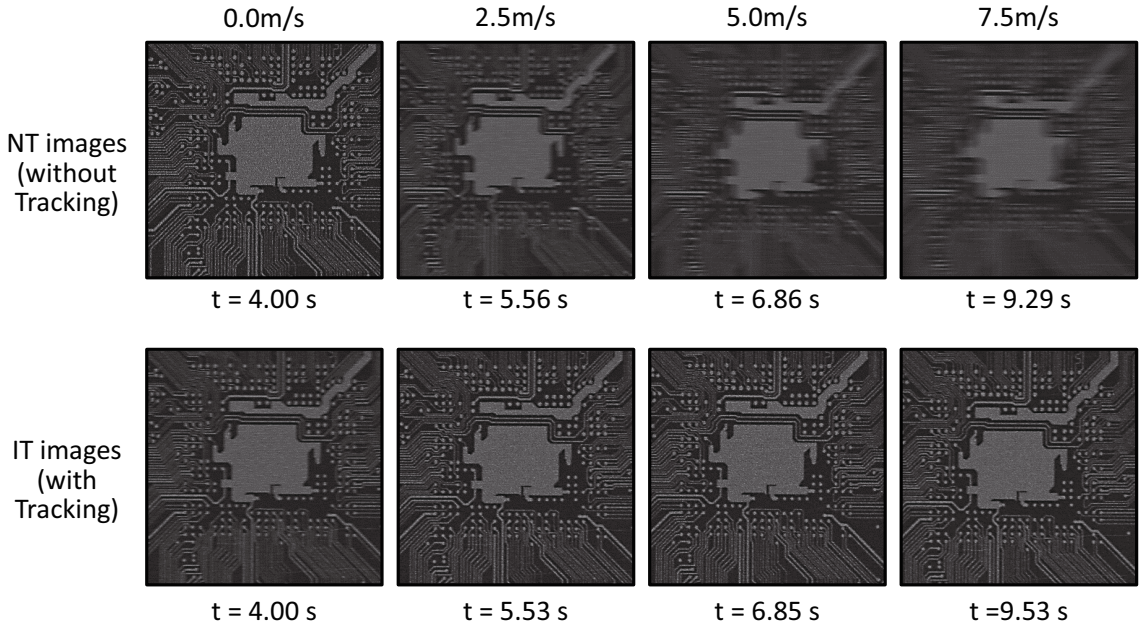


Figure 5.18: Images captured when the electronic board pattern moved at variable speeds.

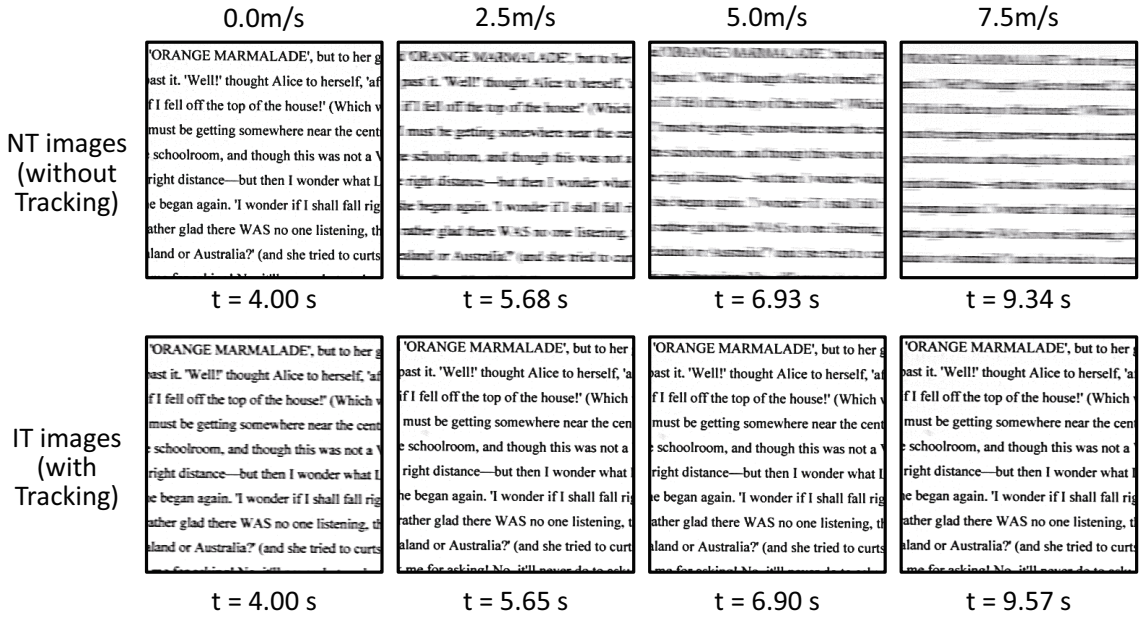


Figure 5.19: Images captured when a book page with many letters moved at variable speeds.

there was no vibration of the resonant mirror. The IT images at all the speeds resembled the images of unmoving patterns, which corresponded to the NT images when the object speed was 0.0 m/s, whereas the motion blur for the NT images became larger in

the horizontal direction as the object speed increased. When the electronic board pattern was moving at 0.0, 2.5, 5.0 and 7.5 m/s, the blur indexes λ_{edge} of the IT images were 11.5, 14.8, 15.7 and 13.4%, respectively, whereas those of the NT images were 19.8, 9.45, 8.13 and 7.68%, respectively. When the book page pattern was moving at 0.0, 2.5, 5,0 and 7.5 m/s, the blur indexes λ_{edge} of the IT images were 5.22, 7.19, 7.59 and 6.76%, respectively, whereas those of the NT images were 8.65, 4.09, 3.62 and 3.51%, respectively. Thus, fast-moving complex patterned objects such as the wiring patterns of 0.25-mm width printed on the electronic board pattern and the 2-mm alphabet letters printed on the book page pattern were observable without noticeable blurring by applying the actuator-driven frame-by-frame intermittent tracking method with amplitude control.

5.5 Conclusion in this Chapter

In this section, we developed a motion-blur-free high-frame-rate video shooting system based on the concept of actuator-driven frame-by-frame intermittent tracking. In this system, the constant camera frame timings are controlled for video shooting with a larger camera exposure time in synchronization with the high-frequency free vibration with a large amplitude of a resonant mirror so that it enables the ultrafast gaze control to track fast-moving objects when the camera shutter is open, whereas the viewpoint velocity is controlled by the mirror amplitude of the resonant mirror. Our system can capture 1024×1024 images of fast-moving objects at 750 fps with an exposure time of 0.33 ms without motion blur, and its performance was verified by conducting several video shooting experiments for fast-moving patterned objects on a high-speed belt-conveyor system.

Chapter 6

Motion-blur-free High-frame-rate Video Camera with Frame-by-Frame Visual-Feedback Control

In this chapter, to verify the effectiveness of the actuator-driven frame-by-frame intermittent tracking and the blur-index-based visual-feedback algorithm, we developed an FPGA-based motion-blur-free HFR video camera system using a resonant mirror with an edge-intensity-based visual-feedback control.

6.1 Camera Shutter Timings and Exposure Times

As mentioned above in the previous chapter, it is important to determine a longer camera exposure time with consideration of the permissible deviation error in straight-line approximation, which corresponds to the degree of motion blur, in motion-blur-free video shooting with actuator-driven frame-by-frame intermittent tracking. The tracking trajectory during the exposure time has the shape of the nonlinear sinusoid trajectory with resonant vibration of a free-vibration-type actuator. In this section, we discuss how to determine parameters for camera shutter timings and exposure time in actuator-driven frame-by-frame intermittent tracking based on the mathematical relationship between the segmented sinusoid trajectory and its approximate straight line.

As illustrated in Figure 6.1, the input image is captured at frame k with an exposure time τ and the variable frame-capture-event (FCE) τ_d by opening and closing the camera

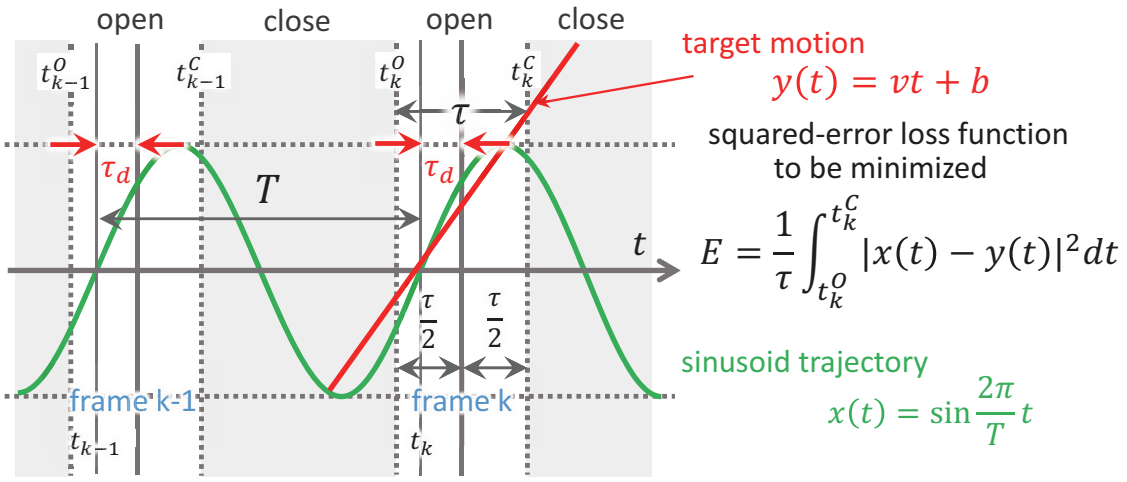


Figure 6.1: Sinusoid trajectory and its approximate straight line.

shutter at times $t_k^O = t_k + \tau_d - \tau/2$ and $t_k^C = t_k + \tau_d + \tau/2$, where $t_k = 2n\pi$ (n: integer) is the vibrating center time of the sinusoid trajectory $x(t) = \sin(2\pi/T)t$ in the each period. To control the camera viewpoint for a target object moving at a speed of v during the camera shutter open, the FCE τ_d is controlled so that the straight line $y(t) = vt + b$ approximates the segmented sinusoid trajectory in the range of time t_k^O to time t_k^C. Here, we assume that the vibrating center time $t_0(k) = 0$ is set for simplification, and consider a minimization problem for the following squared-error loss function E that can evaluate the deviation of the segmented sinusoid trajectory from the straight line where the target object moves,

$$E = \frac{1}{t_k^C - t_k^O}\int_{t_k^O}^{t_k^C} |x(t) - y(t)|^2 dt \tag{6.1}$$

The following relation is satisfied when the squared-error loss function E becomes minimum:

$$\frac{\partial E}{\partial v} = 0, \quad \frac{\partial E}{\partial b} = 0 \tag{6.2}$$

Based on formula 6.1, formula 6.2 can be converted as follows:

$$\begin{cases} v\int_{t_k^O}^{t_k^C} t^2 dt + b\int_{t_k^O}^{t_k^C} t dt - \int_{t_k^O}^{t_k^C} t\sin\frac{2\pi}{T}dt = 0 & (6.3) \\ v\int_{t_k^O}^{t_k^C} t dt + b\int_{t_k^O}^{t_k^C} dt - \int_{t_k^O}^{t_k^C} \sin\frac{2\pi}{T}dt = 0 & (6.4) \end{cases}$$

Moreover, formula 6.2 can be converted as follows:

$$\begin{cases} v\tau(\tau_d^2 + \frac{\tau^2}{12}) + b\tau_d\tau - \frac{T^2}{2\pi^2}\cos 2\pi d \sin\pi r \\ \quad + \frac{T}{2\pi}(\tau\cos 2\pi d\cos\pi r - 2\tau_d \sin 2\pi d \sin\pi r) = 0 & (6.5) \\ v\tau_d\tau + b\tau - \frac{T}{\pi}\sin 2\pi d\sin\pi r = 0 & (6.6) \end{cases}$$

where $r = \tau/T$ and $d = \tau_d/T$ are the time open ratio and the delay time ratio, which indicates the ratio of the exposure time τ and the FCE $\tau_d, \Delta t$ to the cycle time T of the sinusoid trajectory, respectively. The slope v and y-intercept b of the straight line $y(t)$ is shown as follows:

$$v = \frac{3\cos 2\pi d(\sin\pi r - \pi r\cos\pi r)}{\pi^3 r^3}\omega \tag{6.7}$$

$$b = \frac{\pi r^2 \sin 2\pi d\sin\pi r + 6d\cos 2\pi d(\pi r\cos\pi r - \sin\pi r)}{\pi^2 r^3} \tag{6.8}$$

Solving the following equation such that the partial derivative of $E(d)$ for the delay time rate d is zero,

$$\frac{\partial E(d)}{\partial d} = \frac{\sin 4d\pi}{2\pi^3 r^4}(6 - 5\pi^2 r^2 - (6 + \pi^2 r^2)\cos 2\pi r + \pi r(2\pi^2 r^2 - 3)\sin 2\pi r) \tag{6.9}$$

the delay time d_{min}, which set the squared-error loss function E for 0, can be derived as $d_{min} = 0, 0.5, 1 (0 \le d \le 1)$. Figure 6.2 shows the relationship between $\min E(d)$ and

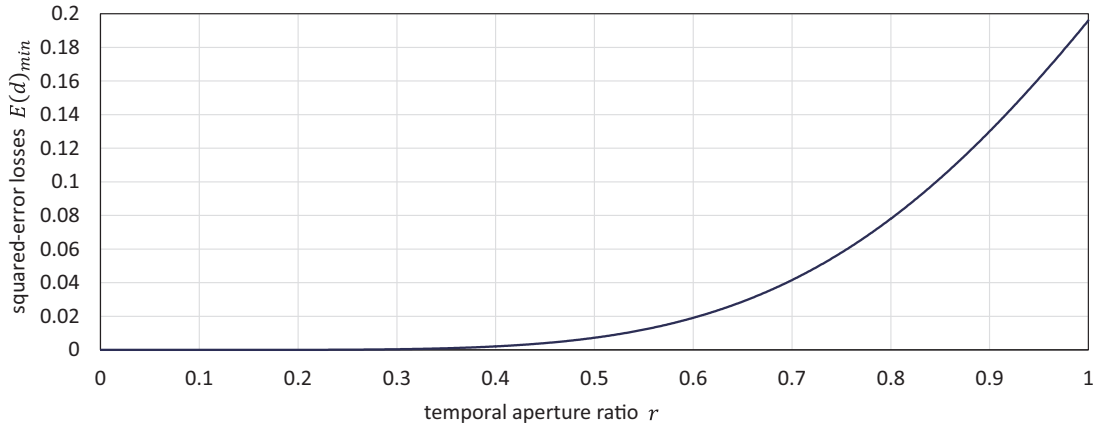


Figure 6.2: Relationship between $\min E(d)$ **and time open ratio** r**.**

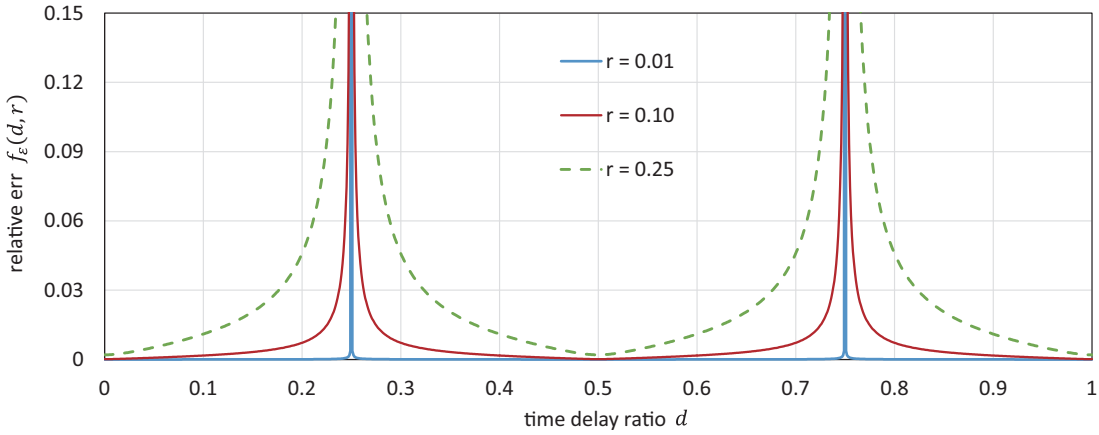


Figure 6.3: Relationship between aperture ratio r**, delay time ratio** d **and relative error ratio** ε**.**

time open ratio r.

The relative error ratio ε per unit the cycle time T is set as follows:

$$\begin{aligned} \varepsilon &= |\frac{\Delta v}{v}| \cdot \frac{1}{T} \\ &= |f_\varepsilon(d,r)| \end{aligned} \tag{6.10}$$

where $\Delta v = \sqrt{E}$ is the displacement in the range of time t_k^O to time t_k^C. Figure 6.3 shows the relative error ratio $|f_\varepsilon(d,r)|$, when $r = 0.01$, $r = 0.10$, $r = 0.25$. Using

the relationship between r, d and ε in Figure 6.3 as a look-up table, we set the camera exposure time τ and the FCE τ_d, depending on the selected relative error ratio value ε.

6.2 Developed Motion-blur-free HFR Video Camera System with Blur-Index-Based Visual-Feedback Control

6.2.1 Outline of System

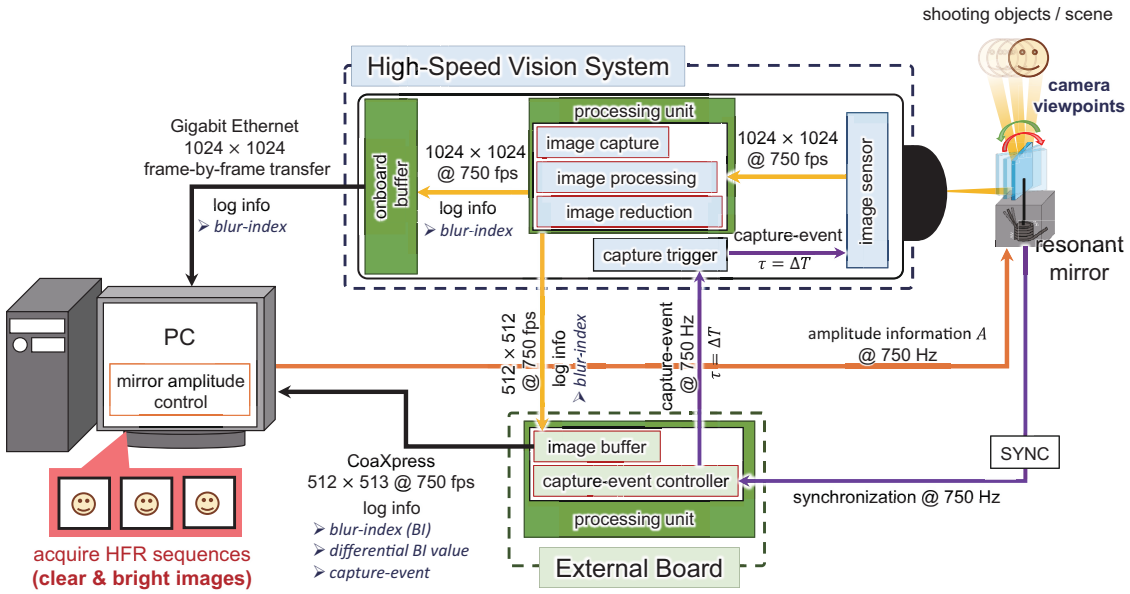


Figure 6.4: Outline of FPGA-based motion-blur-free HFR video camera system using resonant mirror with blur-index-based visual-feedback.

Figure 6.4 shows the outline of proposed system consists of a high-speed vision system for capturing 1024×1024 pixels size HFR image sequences and processing in real-time at device level; an external board for onboard processing of 512×512 pixels images and interfacing with personal computer (PC); the resonant mirror which is mounted in the optical axis of the high-speed vision system for sensing multi-view optical viewpoint velocity and the PC for real-time management of instruction using software coding interface. The high-speed vision system and the external board are equipped with FPGA enabled image processing units. The high-speed vision system is facilitated with onboard memory to buffer the motion-blur-free HFR image sequences to overcome the limitation

of bandwidth of conventional interface between camera and PC. It can transfer buffered motion-blur-free images to PC hard drive using a gigabit ethernet interface. In order to preprocess image sequences using hardware logic and map them to PC memory in real-time, the external board is integrated with the high-speed vision system. It interfaces with PC using CoaXpress high-speed interface.

The high-speed vision system shoots the targeted objects reflected by a resonant mirror by vibrating its mirror at a large amplitude at 750 Hz. Image sequences from both the high-speed vision system and the external board are combined with logged information such as processed output features, frame number, etc. for further feedback-based processing and control. The frame-capture-event (FCE) control unit of the external board controls the image capture event of the high-speed vision system based on a frame-by-frame close-loop visual-feedback calculated using the blur-index-based degree of sharpness when the FCE control. The external board also has onboard memory to buffer motion-blur-free image sequences to which log information is appended at a single bottom row of the image. Hence our system can transfer motion-blur-free images of pixel size 512×513 in real-time to PC memory and 1024×1024 pixels size motion-blur-free images offline using CoaXpress and gigabit ethernet interfaces respectively. When the motion-blur-free HFR video camera system selects the amplitude control for the resonant mirror, the analog information of 0∼5 V, which is calculated by the processing results of differential blur-index value from the external board via the CoaXpress is transferred from a DA board of the PC.

6.2.2 Constructed System Component

We fabricated a motion-blur-free HFR vision system, as shown in Figure 6.5. It shows the overview of the fabricated prototype with mirror units of a high-speed vision system, a planar mirror, and a resonant mirror. We used the similar high-speed conveyor system in Chapter 5 for a verification experiment. The motion-blur-free HFR vision system consists of high-speed vision platform developed by FASTCAM SA-X2 (Photron) with an F-mount 135-mm mounted lens (Rodagon 135 mm F5.6, Qioptiq Photonics), an

external board (Techno-Scope), a resonant mirror (SC-30, Electro-Optical Products), a function generator (AFG1022, Tektronix) and a personal computer (PC) with a supermicro motherboard Super X10SRA, Intel Xeon CPU E5-1630 v3 3.70GHz, 8-GB memory, Windows 7 (Enterprise 64 bit OS), a D/A board (PEX-340416, Interface, Hiroshima, Japan) and a CoaXPress board (APX-3664, Aval Data). The high-speed belt-conveyor in Chapter 5.2 was located with a similar position in figure 5.4 at the shooting distance 760 mm between the resonant mirror and the shooting objects and used for driving the various patterns in shooting experiments.

The high-speed vision system has a 12-bit 1024×1024 CMOS image sensor, the sensor size and pixel size of which are 20.48×20.48 mm and 20×20 μm, respectively, and it can capture and record 1024×1024 images at 12,500 fps in onboard memories with a 256-parallel analog output at a 65-MHz clock. The camera parameters, such as exposure time, can be set through a gigabit ethernet from external systems. In contrast, an external trigger signal can directly determine the image capture event to open the camera exposure. We implemented the image processing units with the image processing in two memories (MEM FPGA) of the camera system.

SC-30 is a resonant mirror that can vibrate a 23 mm×23 mm-size glass mirror in the pan direction at its resonant frequency of 750 Hz. The amplitude of its sinusoid trajectory can be controlled in the range of 0.0025 to 0.5° by providing a voltage command in the range of 0 to 5 V to its automatic gain control driver. A transistor-transistor-logic (TTL) signal is externally output at the cycle time of its resonant vibration, and it was used to determine the FCE to open the camera exposure at 750 Hz.

The function generator AFG1022 was used to adjust the FCE in the external TTL signal to synchronize the center time of the open exposure with the vibration center of the sinusoid trajectory at no FCE control and input the adjusted TTL signal into the D/IO port of the external board.

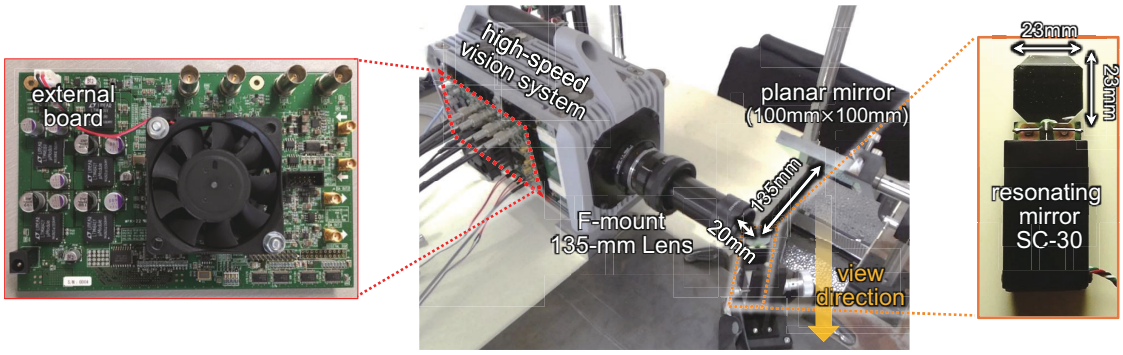


Figure 6.5: An overview of fabricated motion-blur-free HFR vision system.

6.2.3 Implemented Hardware Logic

To execute motion-blur-free frame grabbing in real-time with sub-microseconds precision, we implemented image processing and frame-capture-event (FCE) control logic on the processing units (PU) of the high-speed vision system and the external board. As shown in Figure 6.6 (a) and (b), the high-speed vision system and external boards are equipped with Altera Stratix IV (EP4SGX290KF43C2) FPGA-based PU-1 and PU-2, respectively.

The hardware logic implemented on PU-1 is capable of processing 1024×1024 pixel-sized images in real-time at 12,500 fps [105]. It also synthesizes edge detection as the blur-index and the image reduction process using with 259.2 MHz clock speed of FPGA, and it is enabled with an onboard image buffering unit for high-speed image acquisition and real-time processing. However, the FCE control for triggering capture signals as a visual-feedback control is processed by the external board integrated with the high-speed vision system. It is equipped with PU-2, which is also Altera Stratix IV (EP4SGX290KF43C2) FPGA-based hardware logic implementation platform. PU-2 is interfaced with the resonant mirror to synchronize mirror oscillation frequency, i.e., 750 Hz with frame capture rate 750 fps of the high-speed vision system. PU-1 and PU-2 are interconnected via a high-speed interface for transferring processed image features and information of each pixel. The pixel-thinned images generated by a pixel reduction module in PU-1 of pixel size 512×512 are simultaneously transferred to PU-2 while capturing the 1024×1024 pixels size images. The edge detection submodule calculates the

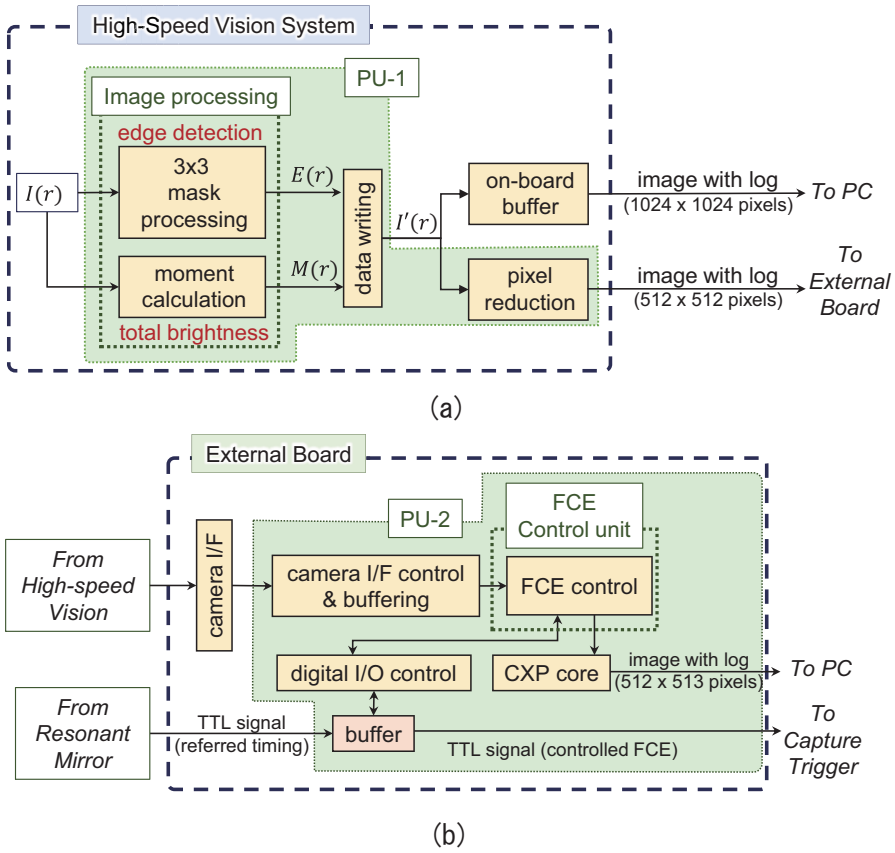


Figure 6.6: (a) High-speed vision on-board PU-1 for hardware logic of edge detection and total brightness, (b) External board with PU-2 for hardware logic of frame-capture-event (FCE) control and trigger.

edge intensities $E(r)$ using a 3×3 Laplacian neighborhood kernels with reference to frame number r. It passes the image processing results to the data output module, whereas, moment calculation submodule simultaneously extracts zeroth moment as an image brightness $M(r)$ regarding frame number r to pass it to the data output module. The PU-2 on external board buffers the thinned pixels and edge gradient features for calculating the normalized edge intensities $R(r)$, the subtraction $C(r)$ between the different speeds of the camera viewpoint and control the FCE τ_d of the high-speed vision system.

Table 6.1 presents the resource consumption when the hardware logic implemented in PU-1 of the high-speed vision system and PU-2 of the external board. Hence, the proposed hardware architecture could be implemented in the available resources of the

Table 6.1: FPGA resource consumption

Device Type	Altera Stratix IV	EP4SGX290KF43C2
	High-speed-vision	External board
LEs	68,269/291,200	25,792/291,200
ALMs	14,053/116,480	1,298/116,480
Registers	61,242/232,950	25,143/232,950
M9K blocks	812/936	934/936
M144K blocks	32/36	28/36
Embedded Memories	8,960/13,608	10,539/13,608
MULT18X18s	576/832	56/832
PLL	9/12	5/12

onboard FPGAs.

The external board consists of one MEM FPGA that incorporate image data into the PU-2. The frame-capture-event-controller-unit (FCECU) of PU-2 executes a frame-by-frame FCE control by inputting a high frame rate image, which was transferred from the 8ch×2 high-speed external LVDS port of the high-speed vision platform. The 12-bit grayscale 512×512 images, which was decimated from 1024×1024 images in the image thinning function of the PU-2 with an additional row of stored log information, were transferred to the PC by CoaXpress I/O port (CXP-6, 6.25 Gb/s, 4 ch). The TTL signal timings are controlled in accordance with image processing results in the FCECU of the MEM FPGA, and output to the high-speed vision system as an external trigger signal through the D/IO port.

6.2.4 Implemented Algorithm

In the actuator-driven frame-by-frame intermittent tracking, the camera viewpoint is controlled at the sinusoid trajectory by a resonant mirror. In this study, when A is the mirror amplitude of the resonant mirror and τ_s is the center time of the camera exposure, the viewpoint's position $p(t)$ at time t indicates at a period time of $T = 1/f_0$ as the following sinusoidal trajectory:

$$p(t) = A(t) \cdot \sin \frac{2\pi}{T} t. \tag{6.11}$$

6.2.4.1 Preliminary setting

For shooting the images at two little different viewpoint speeds, the variable frame-capture-event of the high-speed vision system is controlled alternately in even and odd frames. The input image is captured at frame k with exposure time. The center exposure time τ_s is controlled on the following timings (n: integer):

$$\tau_s = \begin{cases} t_k + \tau_d(k), & (k = 2n+1) \\ t_k + \tau_d(k) + \Delta t. & (k = 2n) \end{cases} \tag{6.12}$$

where $t_k = 2n\pi$ is the center time of the sinusoid trajectory $p(t)$ in each period, and τ_d can control the viewpoint speed during the exposure time in the total frames ($0 \leq \tau_d \leq T$). Δt in even frames is set for capturing the different events alternately to control the viewpoint-velocity of the camera frame-by-frame [106].

6.2.4.2 Real-time processing

For realizing the visual-feedback control for the camera viewpoint velocity in the frame-capture-event control and the amplitude control, respectively, the following algorithm was implemented on the prototype system.

(1) Calculation of brightness/edge intensity

A gray-scale 1024×1024 input image $I(k)$ is captured at frame k. The total brightness $M(k)$ and the edge intensity $E(k)$ are calculated from $I(k)$.

(2) Calculation of normalized differential edge intensity

The normalized difference of edge intensities (DEI) $C(k)$ as a DBI is calculated from the normalized edge intensities $R(k) = E(k)/M(k)$ as follows (n:integer):

$$C(k) = \begin{cases} -R(k) + R(k-1), & (k = 2n+1) \\ R(k) - R(k-1). & (k = 2n) \end{cases} \tag{6.13}$$

(3) Frame-capture-event control or amplitude control based on DEI

Considering K_{dp} and K_{da} are the proportional gains, the frame-capture-event τ_d and

the mirror amplitude A are controlled with $C(k)$ as follows, respectively:

$$\tau_d(k) = \tau_d(k-1) + K_{dp} \cdot C(k). \tag{6.14}$$

$$A(k) = A(k-1) - K_{da} \cdot C(k). \tag{6.15}$$

6.2.5 Parameter setting

With this vision system, the 1024×1024 input images were captured at 750 fps, with the exposure time of 0.30 ms (= 1/3333) and the temporal time aperture ratio of $r = 0.225$ in synchronization with the external trigger signal as the controlled frame interval from the resonant mirror via the external board. A shooting region of 1024×1024 pixels corresponded to the 104×104 mm shooting area of a constant shooting distant 760 mm, and 0.102 mm corresponded to one pixel. The frame-capture-event Δt in even frames was set to 0.07 ms, whereas the frame-capture-event τ_d can be controlled in the range of $\tau_d = 0.00$ ms to $\tau_d = 0.30$ ms. Table 6.2 shows the relationship between a relative error ε and a controlled frame-capture-event $\tau_D (= \tau_d + \Delta t)$ ms; this corresponded to the relationship between r, d, and ε in Figure 6.3.

The frame-capture-event control and the amplitude control can be selected before experimenting. The mirror amplitude A was set as a constant in the frame-capture-event control, whereas the frame-capture-event τ_d was set to zero in the amplitude control. The video camera system started the either of controls when the total brightness $M(k)$ of the captured image exceeded the threshold 2.5×10^8. After passing the shooting objects and

Table 6.2: Relationship between a relative error ε and a controlled frame-capture-event $\tau_D (= \tau_d + \Delta t)$ ms

controlled frame-capture-event τ_D	relative error ε
0.00 ms	0.14×10^{-2}
0.05 ms	0.32×10^{-2}
0.07 ms	0.44×10^{-2}
0.10 ms	0.63×10^{-2}
0.15 ms	1.04×10^{-2}
0.20 ms	1.66×10^{-2}
0.25 ms	2.91×10^{-2}
0.30 ms	7.60×10^{-2}

getting dark in images, the controller set $\tau_d(k)$ to zero and stored $A(k)$ for the next objects.

6.3 Preliminary Experiments

6.3.1 Displacement of resonant mirror

Firstly, we conducted a preliminary experiment to verify the angular displacement of the resonant mirror. To measure the angular displacement, we irradiated the center of the captured area with a laser beam spot, and the locations of the beam spots via the resonant mirror were identified in offline by capturing an HFR video of 128×48 pixels at 100,000 fps with the exposure time of 8.98 $\times 10^{-3}$ ms. Figure 6.7 shows the angular displacement for 4.0 ms when the drive voltages to the resonant mirror were set to 0.0, 2.5, and 5.0V. Every angular displacement shows a substantially sinusoidal waveform with stability at a frequency of 750 Hz. The averaged vibration amplitudes in 0.0, 2.5 and 5.0V were, 0.0152°, 0.0826° and 0.1634°, respectively. The standard deviations of the mirror vibration in 0.0, 2.5 and 5.0 V were, 2.17 $\times 10^{-4\circ}$, 2.26 $\times 10^{-4\circ}$ and 2.12 $\times 10^{-4\circ}$. Figure 6.8 shows the relationship among the input voltages for the resonant mirror, the mirror displacements of the resonant mirror, and its standard deviations. As the input voltage increased, the mirror displacement of the resonant mirror became large; however, the standard deviations showing its stability maintained. There was a slight offset around 0 V; the relationship between the drive voltage V (V) and the vibration amplitude A (deg) can be linearly approximated as $A = 0.031V + 0.0066$. Figure 6.9 shows the tracking speed v and the relative error ratio ε in $0 \leq \tau_d \leq T/4$ when the drive voltages to the resonant mirror were set to 0.0, 2.5, and 5.0V. As can be seen from this figure, the viewpoint velocity v increased as the drive voltage became progressively larger, and the frame-capture-event τ_d grew large as the relative error ratio ε got gradually wider.

In the frame-capture-event control at several experiments, the mirror displacement of the resonant mirror was set to $A = 0.0692°$ by considering the range of the object speed as the amplitude of an appropriate parameter; the maximum angular speed was $2.07 \times 10^{\circ}/s$ at $A = 0.0692°$. In the amplitude control, the frame-capture-event τ_d was set to 0.00 ms, and the amplitude of resonant mirror was control in the 0.0V to 5.0V range for

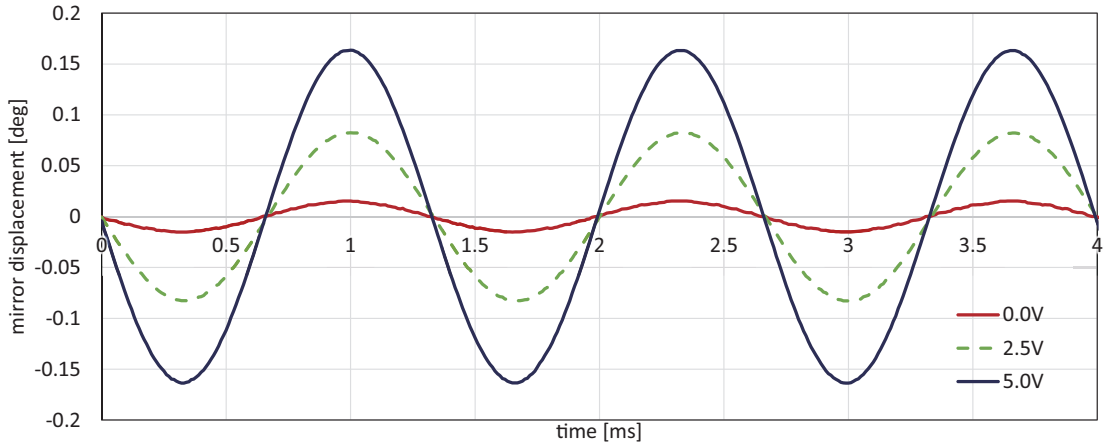


Figure 6.7: Mirror displacement in the drive voltage of 0.0V, 2.5V and 5.0V.

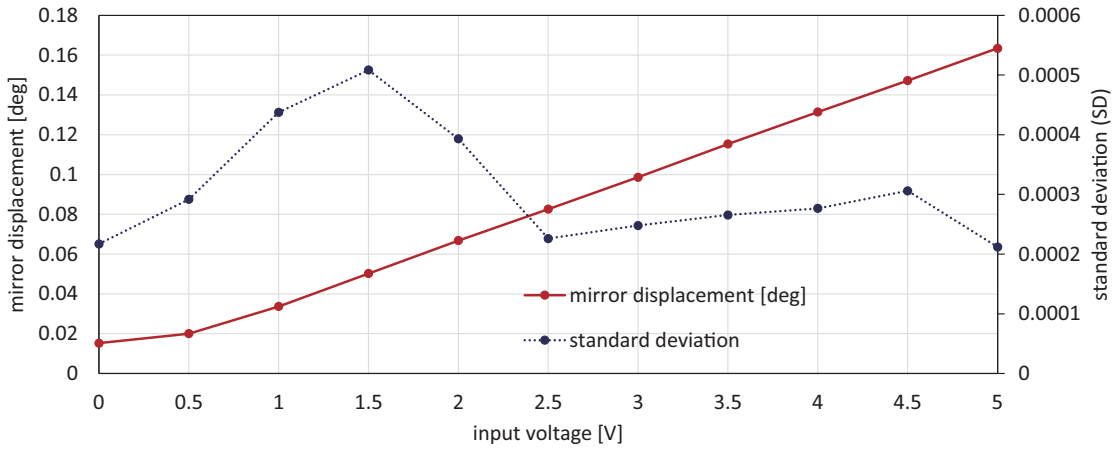


Figure 6.8: Mirror displacement and standard deviation in each drive voltage.

selecting an appropriate viewpoint speed; the maximum angular speed was $4.90 \times 10^{\circ}/s$ at $A = 0.1634^{\circ}$.

6.3.2 Setting of P controller

Next, for determining appropriate parameters of the gain K_{dp} and K_{ap}, we experimented with the continuous cycling method in the Ziegler-Nichols ultimate sensitivity methods [107–109]. A checkered pattern was moved at constant speeds 3.83 m/s using a high-speed conveyor belt system to measure a gain K_u, which caused a continuous oscillation on the controlled value, and we confirmed the gain K_u by increasing the pro-

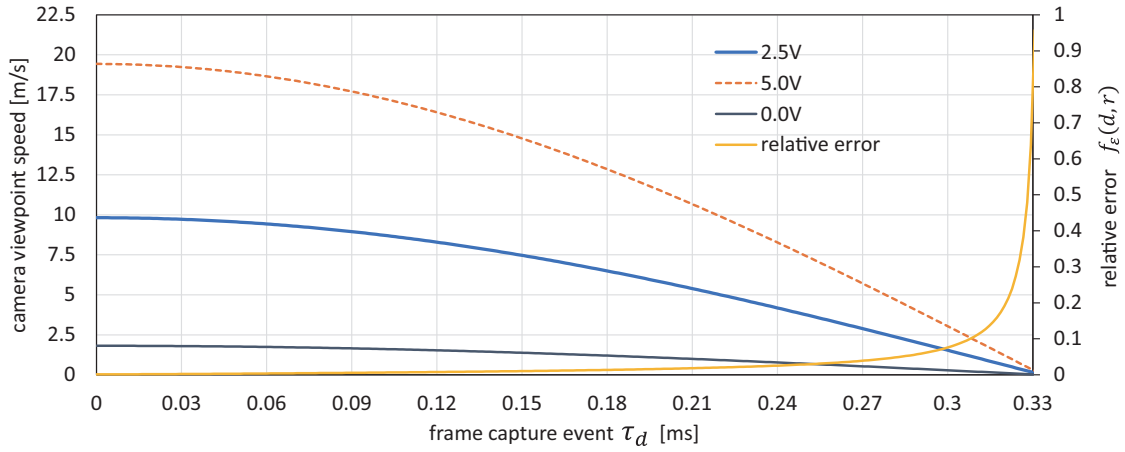


Figure 6.9: Viewpoint velocity in the drive voltage of 0.0, 2.5 and 5.0V and the relative error ε when the time open ration r is 0.225 at shooting distance 760 mm.

portional gain in steps of 0.005. Figure 6.10 shows the patterned object to be observed in a 1024×1024 image. The images feature a pattern of the checkered pattern on which 2×2 mm squares of alternating black and white were printed. Through this preliminary experiment, the parameter of the gains K_{dp} and K_{ap} were 1.190 and 0.0275, respectively, by calculating as $K_p = 0.50K_u$.

6.4 Video Shooting Experiments

6.4.1 Video Shooting without Feedback Control for Checkered Patterns Moving at Constant Speeds

We conducted an experiment to verify the relationship between a controlled frame-capture-event and its sharpness. This video shooting experiment was conducted for a checkered-pattern in Figure 6.10 when the frame-capture-event was set to constant values τ_d = 0.00, 0.05, 0.10, 0.15, 0.20, 0.25 and 0.30 ms (Δt = 0.00). In this experiment, the checkered pattern was moved at constant speeds ranging from 0 to 7.5 m/s at intervals of 0.5 m/s using the high-speed conveyor belt system.

Figure 6.11 shows the 108×108 pixels images that were cropped from the 1024×1024 input images of the checkered pattern moving at motor command of 0.0, 1.0, 3.0, 4.5, 6.5

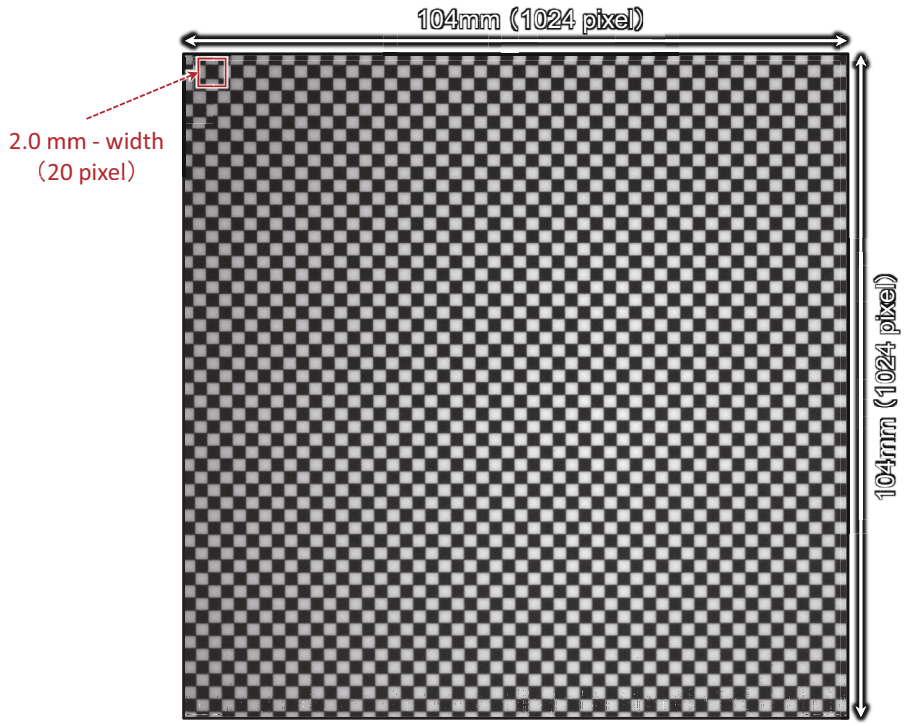


Figure 6.10: Checkered pattern to be observed at 1024×1024 pixels image.

and 7.0 m/s. The figure shows the input images captured without vibration of the resonant mirror (NT, without tracking), which are compared with the input images captured with frame-by-frame intermittent tracking (IT, with tracking) when the frame-capture-event τ_d was set to 0.30, 0.25, 0.20, 0.15 and 0.10 ms. As seen in Figure 6.11, the NT images became increasingly blurred in the horizontal direction as the object speed increased, whereas the IT images at the motor command of 0.0, 1.0, 3.0, 4.5, 6.5 and 7.0 m/s were captured without motion blur when the frame-capture-events were 0.30, 0.25, 0.20, 0.15, and 0.10 ms, respectively. This tendency corresponded to that the squared-error loss functions when the frame-capture-events 0.30, 0.25, 0.20, 0.15, and 0.10 ms were minimized to observe a target object moving at motor commands of 0.0, 1.0, 3.0, 4.5, 6.5 and 7.0 m/s, respectively. The image degradation with motion blur in the horizontal direction increased as the object speed deviated from the desired speed for motion blur reduction, which was determined by the frame-capture-events of the high-speed vision.

For evaluating its blur, the averaged normalized edge intensity R_{ave} was calculated

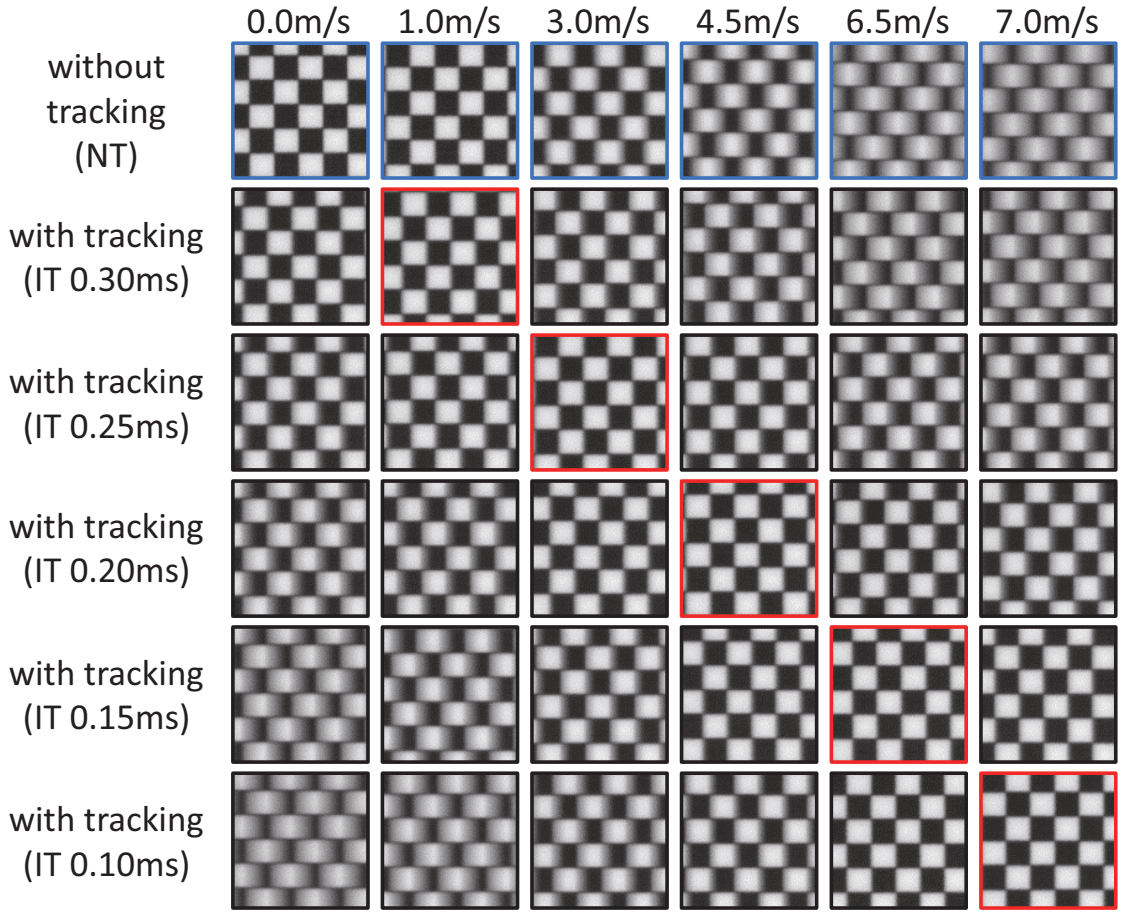


Figure 6.11: Cropped images of a checkered pattern when video shooting without feedback control.

by averaging the normalized edge intensity values of 25 frames. Figure 6.12 shows the relationship between the speed of a checkered pattern and its averaged normalized edge intensity R_{ave} when the target objects moved with the motor command to the conveyor system in the range of 0.0 to 7.5 m/s at intervals of 0.5 m/s. When the frame-capture-event was set to 0.30, 0.25, 0.20, 0.15, 0.10, 0.05 and 0.00 ms in IT images, the averaged normalized edge intensity R_{ave} had a maximum value when the motor commands were 1.0, 3.0, 4.5, 6.5, 7.0, 7.5 and 7.5 m/s, respectively; these motor commands were around the desired speeds for motion blur reduction, 1.29, 3.13, 4.84, 6.24, 7.33, 8.00, and 8.23 m/s, which were determined by the frame-capture-event at 0.30, 0.25, 0.20, 0.15, 0.10, 0.05 and 0.00 ms, respectively. The maximum averaged normalized edge intensities R_{ave} in each frame-capture-event τ_d at 0.30, 0.25, 0.20, 0.15, 0.10, were 0.133, 0.134, 0.138,

0.140, and 0.151 %. Meanwhile, the averaged normalized edge intensities R_{ave} without tracking when the motor commands were 0.0, 1.0, 3.0, 4.5, 6.5 and 7.0 m/s were 0.168, 0.147, 0.118, 0.109, 0.105 and 0.102 %, respectively. As the frame-capture-event τ_d increased, there was a gradual decrease in the averaged normalized edge intensity R_{ave} with tracking when it was the maximum value at each frame-capture-event. This is because of the large error between the linear trajectory of the objects to be shot and the sinusoidal trajectory of the camera's viewpoints.

Through the above experiment results and considerations, the proposed system can control the viewpoint speed for tracking by the frame-capture-event Δt and τ_d. It reduces motion blur in video shooting when the high-speed object corresponds to the desired speed for the motion blur reduction in the frame-capture-event control.

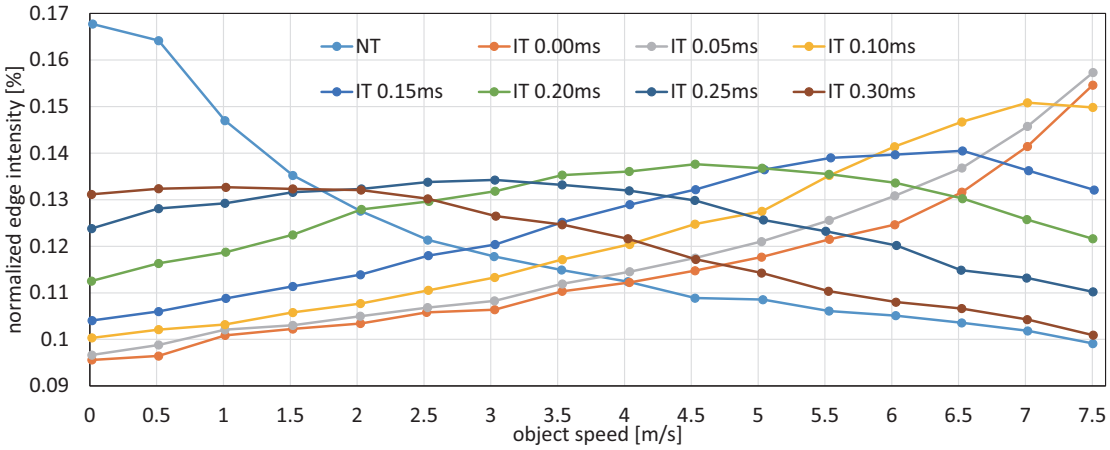


Figure 6.12: Averaged normalized edge intensity R_{ave} when video shooting without feedback control.

6.4.2 Video Shooting with Feedback Control for Checkered Patterns Moving at Constant Speeds

Next, we conducted an experiment for a constant fast-moving patterned object when the frame-capture-event of the high-speed-vision and the amplitude of the resonant mirror were controlled in proportion to the normalized edge intensity difference value C. The checkered pattern that used in the section 6.3.2 was observed in this experiment. Figure

6.13 shows the 108×108 images cropped from the 1024×1024 captured images of the checkered pattern when the patterned objects moved at 0.0, 1.5, 3.0, 4.5, 6.0 and 7.5 m/s by the conveyor system. Figure 6.13 shows the captured images without the vibration of the resonant mirror (NT, without tracking), which are compared with the captured images with frame-by-frame intermittent tracking by the frame-capture-event (IT-FC, with tracking by frame-capture-event) and by the amplitude (IT-AC, with tracking by amplitude). As observed in Figure 6.13, the blur of NT images became larger the object speed in the horizontal direction, whereas the blur of IT-FC and IT-AC images remain small.

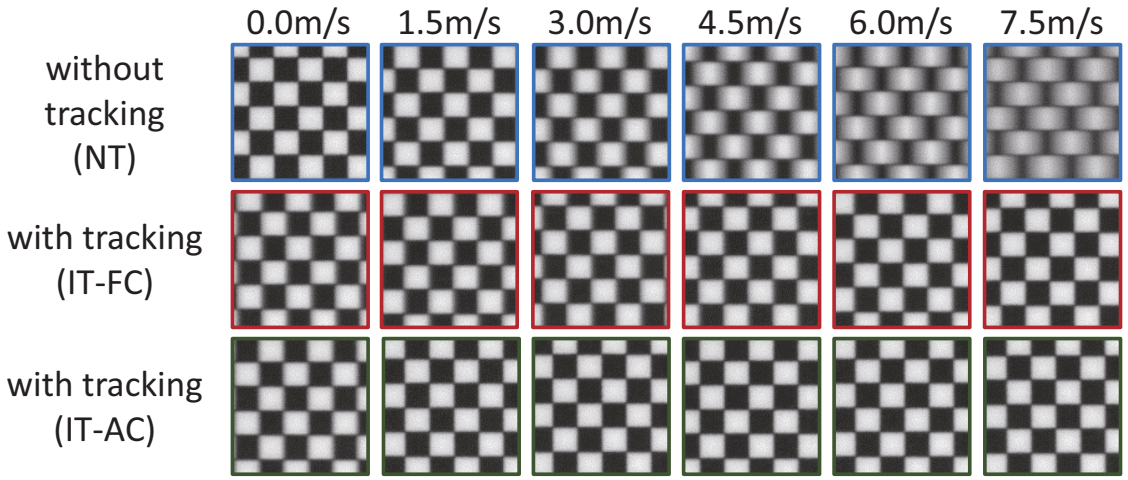


Figure 6.13: Cropped images of a checkered pattern when video shooting with feedback control at constant speed.

Figure 6.14 shows the relationship between the speed of patterns and its averaged normalized edge intensities R_{ave} in the IT-FC, IT-AC and NT images when the conveyor line system was activated in the range of 0 to 7.5m/s at intervals of 0.5m/s. In the figure, the frame-capture-event of the high-speed vision in video shooting the IT-FC images with frame-capture-event control, and the vibration amplitudes of the resonant mirror in video shooting the IT-AC images with amplitude control were also plotted. The index R_{ave} was calculated similarly as in the previous section. Figure 6.15 shows the frame-capture-event τ_d of the IT-FC images with frame-capture-event control and the vibration amplitudes A of the IT-AC images with amplitude control in the range of 0 to 7.5m/s. As can be seen from the Figure 6.15, the frame-capture-events τ_d and the mirror amplitudes A were controlled

for motion blur reduction in proportion to the object speed. There is the vibration of the resonant mirror at a slower speed less than 1.5 m/s in the amplitude control because the mirror amplitude at 0.0152 deg was the minimum value in the resonant mirror. In the Figure 6.14, although the averaged normalized edge intensities R_{ave} for the IT-FC were smaller at a slower speed, the values for the IT-AC images at the faster speeds more than 0.5 m/s was relatively similar with NT images during rest, compared with that for the IT-FC and NT images. This is because the error between the linear trajectory of objects and the sinusoid trajectory of the resonant mirror remained small, comparing with the frame-capture-event control. The averaged normalized edge intensity R_{ave} for the IT-AC images when the speed of the conveyor system was set to 0.0, 1.5, 3.0, 4.5, 6.0 and 7.5 m/s was 0.131, 0.163, 0.162, 0.161, 0.159 and 0.155 %, respectively, whereas that for the IT-FC images was 0.131, 0.130, 0.133, 0.136, 0.141 and 0.152 %, respectively, and that for the NT images was 0.168, 0.135, 0.118, 0.109, 0.105 and 0.099 %, respectively.

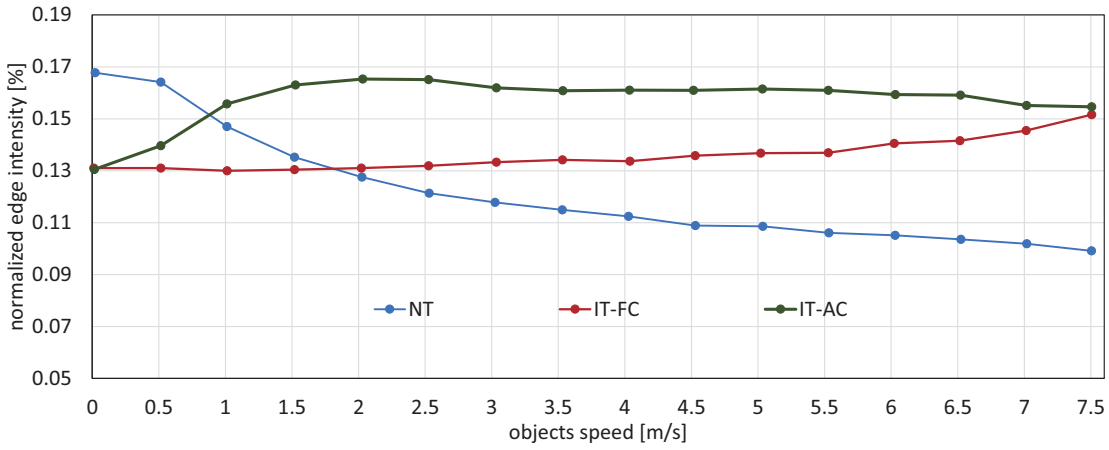


Figure 6.14: Normalized edge intensity R_{ave} when video shooting with feedback control at constant speed.

Through this experiment, it was found that the amplitude control and the frame-capture-event control were controlled in proportion to the speed of the shooting objects, and especially when the amplitude control was useful for the objects at constant speeds.

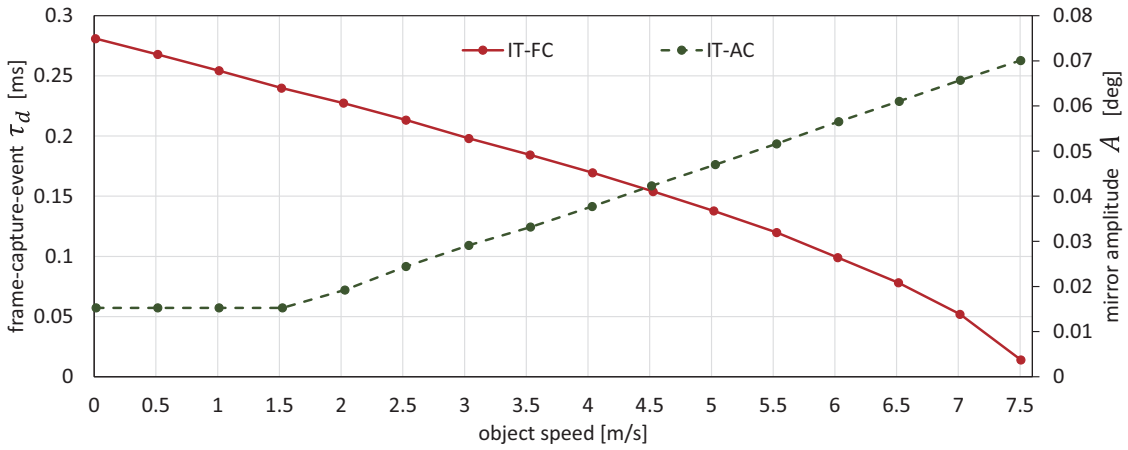


Figure 6.15: Frame-capture-event τ_d and mirror amplitude A_k when video shooting with feedback control at constant speed.

6.4.3 Video Shooting with Feedback Control for Patterned Objects at Variable Speeds

Next, we conducted the video shooting experiments for variable fast-moving patterned objects as the unstable fast-moving scene. In this experiment, the speed of the conveyor system printed the checkered pattern of Figure 6.10 was accelerated from 0.0 m/s to 7.5 m/s in the period of T_{rising} s. The period of T_{rising} was set to 1.3, 3.0, 5.0, 7.0, 9.0, 11, 13, 15, 17, 19, 21, 23, and 25 s. During the accelerating time from $t = 0.0$ s to $t = T_{rising}$ s, the speed of the conveyor system accelerated linearly.

Figure 6.16 shows the result of the normalized edge intensities R and the conveyor speed in NT, IT-FC and IT-AC at the accelerated time $T_{rising} = 25$ s. Figure 6.17 shows the result of the frame-capture-events τ_d in IT-FC and the mirror displacement amplitudes A in IT-AC. In the Figure 6.16, the normalized edge intensities R in NT decreased monotonically in proportion as the conveyor speed, and the normalized edge intensities R in IT-FC and IT-AC increased conversely. As seen in Figure 6.16, the normalized edge intensities R in IT-FC increased monotonically, however, after the normalized edge intensities R in IT-AC decreased once in the middle of increasing, and increased until $t = 25$ s. This is because the controlled mirror amplitudes A were constant until the conveyor speed was approximately 1.5 m/s, and the normalized edge intensities R decreased in the period from

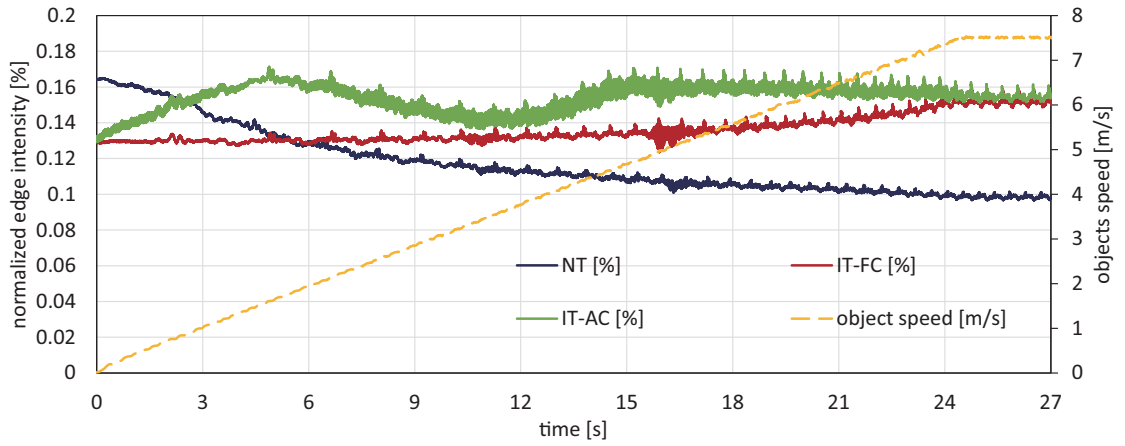


Figure 6.16: Normalized edge intensity R and conveyor speed at T_{rising} =25 s.

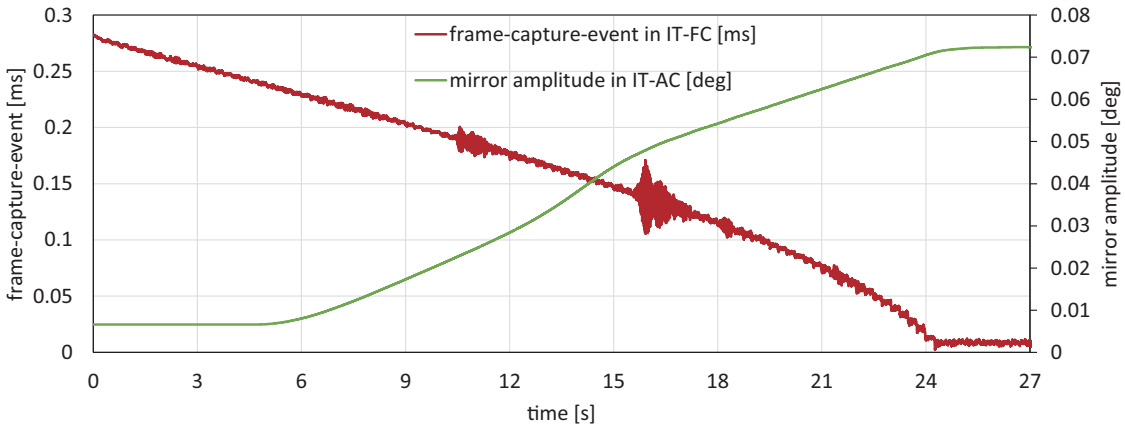


Figure 6.17: Frame-capture-event τ_d and mirror amplitude A_k at T_{rising} =25 s.

5.0 to 12.1 s and tracked the conveyor speed after 12.1 s. In the Figure 6.17, the mirror amplitudes A were controlled according to the conveyor speed, relatively smoothly. In contrast, the frame-capture-event τ_d greatly fluctuated in the period from 15.7 to 17.3 s because the frame-capture-event control with high responsibility was controlled by a jagged joint of the belt which has a locally large value of the edge intensity.

Figure 6.18 shows the result of the normalized edge intensities R and the conveyor speed in NT, IT-FC and IT-AC at the accelerated time T_{rising} = 1.3 s. Figure 6.19 shows the result of the frame-capture-events τ_d in IT-FC and the mirror displacement amplitudes A in IT-AC at T_{rising} = 1.3 s. In the Figure 6.18, the normalized edge intensities R in NT de-

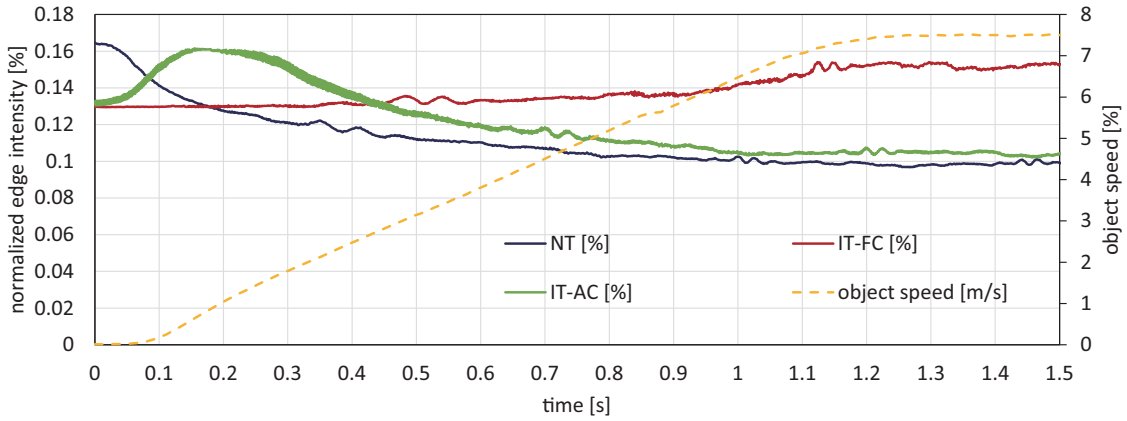


Figure 6.18: Normalized edge intensity R and conveyor speed at T_{rising} =1.3 s.

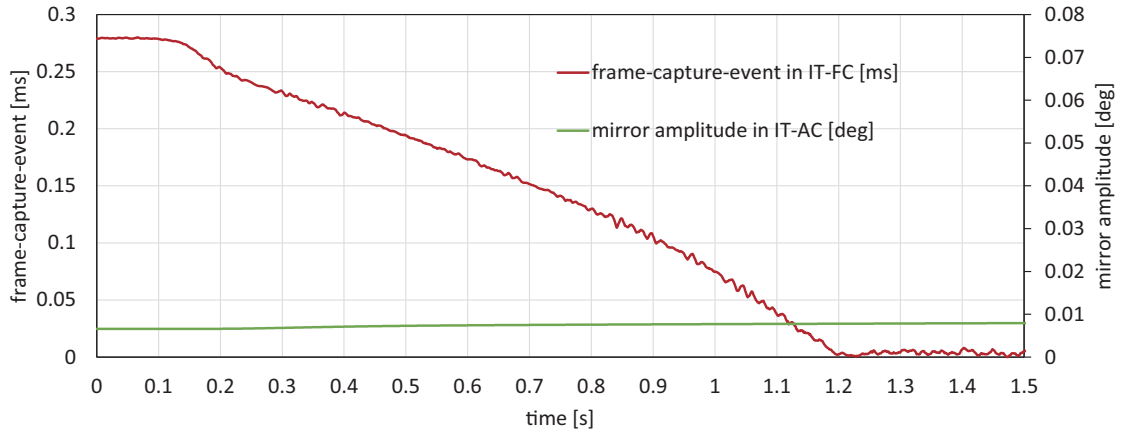


Figure 6.19: Frame-capture-event τ_d and mirror amplitude A_k at T_{rising} =1.3 s.

creased instantly in proportion as the conveyor speed, and the normalized edge intensities R in IT-FC increased because the relative-error decreased and the frame-capture-events τ_d in IT-FC increased in proportion as the conveyor speed. In the IT-AC, the normalized edge intensities R in IT-AC decreased after it increased once. This is because the camera system in IT-AC cannot track beyond the range of its controllability. In the Figure 6.19, the frame-capture-event τ_d did not have a fluctuation compared with T_{rising} = 25 s because the joint of the belt passed, immediately.

Figure 6.20 shows the normalized edge intensity R at $t = T_{rising}$ in NT, IT-FC and IT-AC. As seen in Figure 6.20, the normalized edge intensity R at $t = T_{rising}$ in NT were

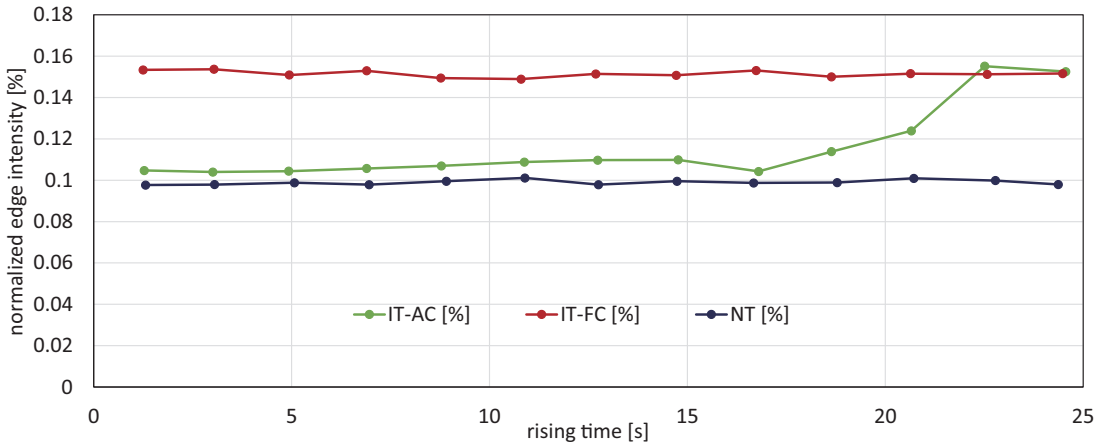


Figure 6.20: Normalized edge intensity R at $t = T_{rising}$.

small because of motion-blur at the conveyor speed 7.5 m/s. In the IT-AC, the camera system can track at T_{rising} = 23 s and 25 s, however, the normalized edge intensities R at $t = T_{rising}$ were small when T_{rising} is less than 23 s because the system in IT-AC cannot track the acceleration of the conveyor speeds below 0.33 m/s^2. On the other hand, in the IT-FC, the normalized edge intensities R was approximately constant even when the accelerating times at T_{rising} were set in the range from 25 s to 1.3 s. These results show that the system in IT-FC can track at 6.79 m/s^2 or more. The acceleration value 6.79 m/s^2 is at T_{rising} = 1.3 s.

From the above, the camera system in the IT-AC is better as compared with the IT-FC when the object speed is constant without acceleration because the relative error between the object trajectory and the mirror trajectory is small. However, when the object speed is variable with strong momentum, the camera system in the IT-FC is better as compared with the IT-AC. This is because the system responsiveness of the frame-capture-event control is high owing to the electronic-signal-based control.

6.4.4 Video Shooting with Feedback Control for Complex Objects at Constant Speeds

Next, we conducted a video shooting experiment for complicated pattern objects using the high-speed conveyor system to verify the effectiveness of the proposed concept

which can capture without relying on the object patterns. Figure 6.21, Figure 6.22 and Figure 6.23 show captured complicated pattern objects in a 1024×1024 image when the conveyor system was stopped. All captured complicated patterns have (a) an electrical board pattern with the size of 160×367 mm, (b) a concrete wall pattern with the size of 160×393 mm, and (c) a text pattern with the size of 135×337 mm. As seen in these figures, the electrical board pattern has several wiring patterns, the concrete wall pattern has cracks and holes, and the text pattern has punctuation marks, capital, and lower-case letters.

We compared our method with two motion-deblurring approaches on captured images as a related method of the motion blur. These were the non-blind convolution motion-deblurring (NBC-MD) method and the learning-based blind convolution motion-deblurring (LBC-MD) method. The method of Levin *et al.* [110] is selected as the NBC-

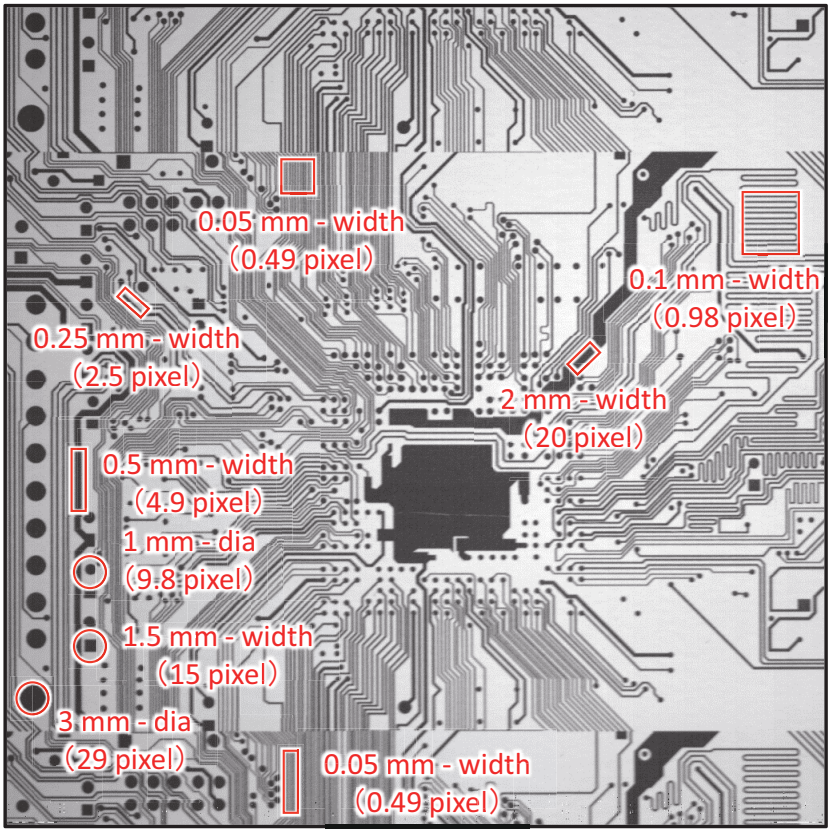


Figure 6.21: Complicated pattern to be observed at 1024×1024 pixels image: (a) Electronic board pattern.

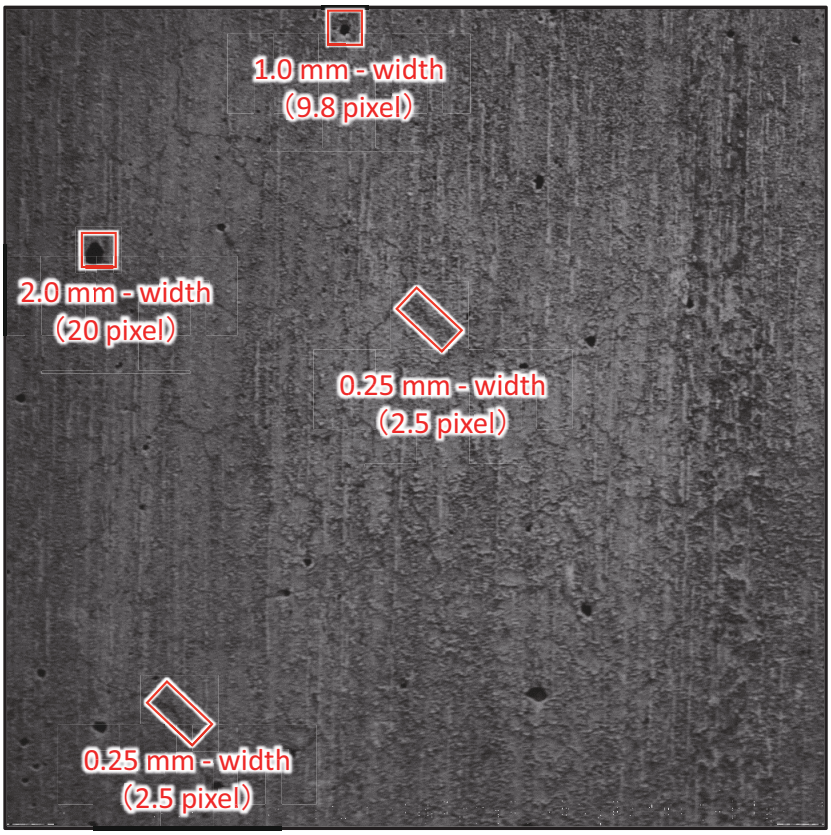


Figure 6.22: Complicated pattern to be observed at 1024×1024 pixels image: (b) Concrete wall pattern.

MD method with a line kernel function, and the method of Tao *et al.* [111] is chosen as LBC-MD method with a scale-recurrent network structure.

Figure 6.24 shows the results of 161×161-pixels-cropped-image sequences in NT, IT-FC and IT-AC when objects moved at 7.5 m/s. Figure 6.24 shows the controlled values of the frame-capture-events and the mirror amplitudes at each captured time as well. The image sequence of NT has substantial motion blur at all times, but the motion blur of the IT-FC and IT-AC image sequence is negligible. There is no change in the mirror amplitude in IT-AC images compared with the frame-capture-event in IT-FC images because the controllability of the viewpoint-velocity control in the amplitude control was lower than the frame-capture-event control.

Figure 6.25(a) shows the 162×162 images of the electrical board pattern, and Fig-

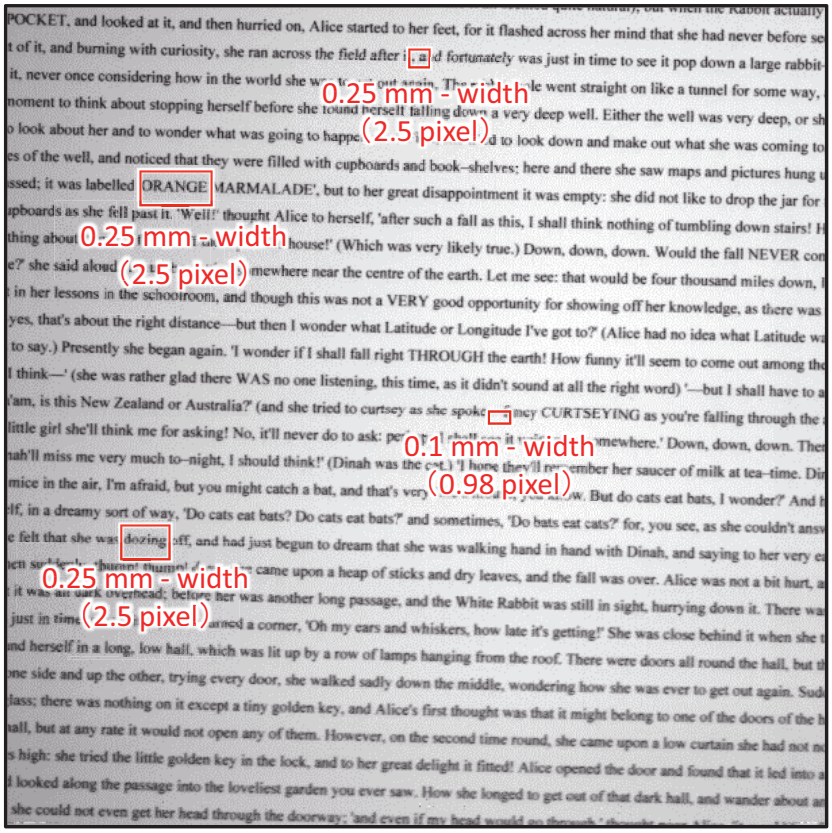


Figure 6.23: Complicated pattern to be observed at 1024×1024 pixels image: (c) Text pattern.

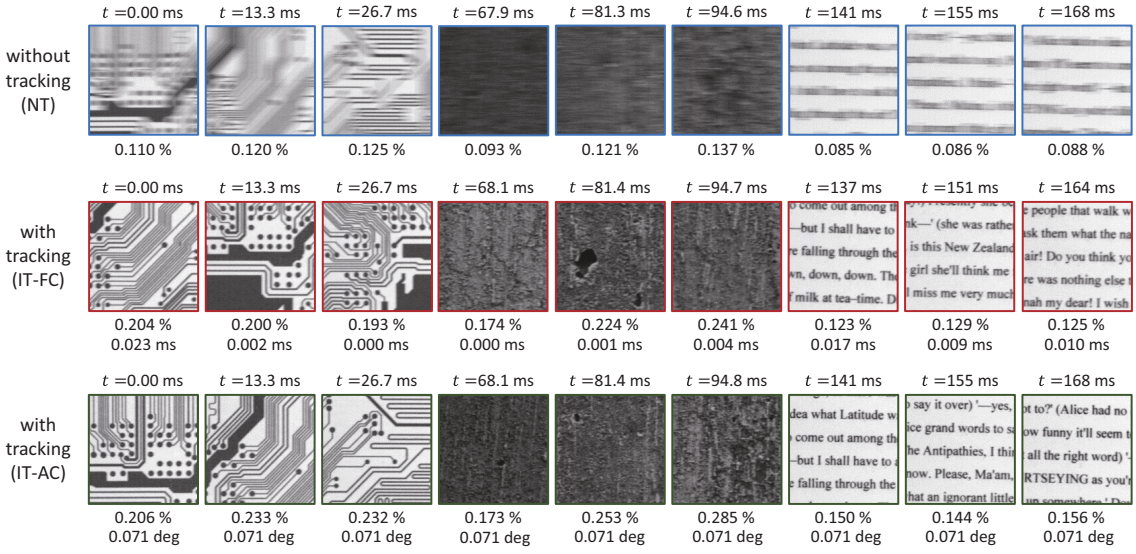


Figure 6.24: Sequence of 161×161-pixel images of NT, IT-FC and IT-AC at 7.5 m/s.

ure 6.25(b) shows the 215×215 images of the concrete wall pattern, and Figure 6.25(c) shows the 162×162 images of the text pattern were cropped from the 1024×1024 images, at 0.0, 1.5, 3.0, 4.5, 6.0 and 7.5 m/s. As seen in Figure 6.25, NT images blurred in the horizon direction according to the conveyor speed. In contrast, IT-FC and IT-AC images did not have observable motion-blurs when the conveyor speed increased even when approximately 22-pixel-blurs in the NT image occur at 7.5 m/s. These results show that the blur-index-based visual-feedback algorithm can control without depending on the object patterns.

Figure 6.26 shows the results of comparison images at each pattern when the speed of the high-speed conveyor was set to 7.5 m/s. As similar to before, the NT image at 7.5 m/s has a large blur compared with the image at the conveyor stop. In the NBC-MD and the LBC-MD images, the motion-blur is significantly less than in the NT images. However, some visible blur remains due to the limitation of the image restoration process when the input image has a remarkable blur. The normalized edge intensity in the center 720×720 pixels of the electrical board pattern 1024×1024-pixels-images in Figure 6.26(a) is 0.129 % in the NT image, 0.150 % in the LBD-MD image, 0.201 % in the NBD-MD image, 0.234 % in the IT-FC image and 0.258 % in the IT-AC images. The normalized edge intensity in the center 720×720 pixels of the concrete wall pattern images in Figure 6.26(b) is 0.130 % in the NT image, 0.120 % in the LBD-MD image, 0.256 % in the NBD-MD image, 0.245 % in the IT-FC image and 0.259 % in the IT-AC images. The normalized edge intensity in the center 720×720 pixels of the text pattern images in Figure 6.26(c) is 0.090 % in the NT image, 0.086 % in the LBD-MD image, 0.155 % in the NBD-MD image, 0.144 % in the IT-FC image and 0.160 % in the IT-AC images. The normalized edge intensities in the NBD-MD images are overall high relative because of the ghost error in the NBD-MD images that is recognized as an image gradient.

The above results indicate the effectiveness of our blur-index-based visual-feedback algorithm using the edge intensity for controlling the viewpoint-velocity of moving objects independently with the patterns of shooting objects. The developed vision system can capture the complicated designs which have the wiring patterns of less than 0.25-mm

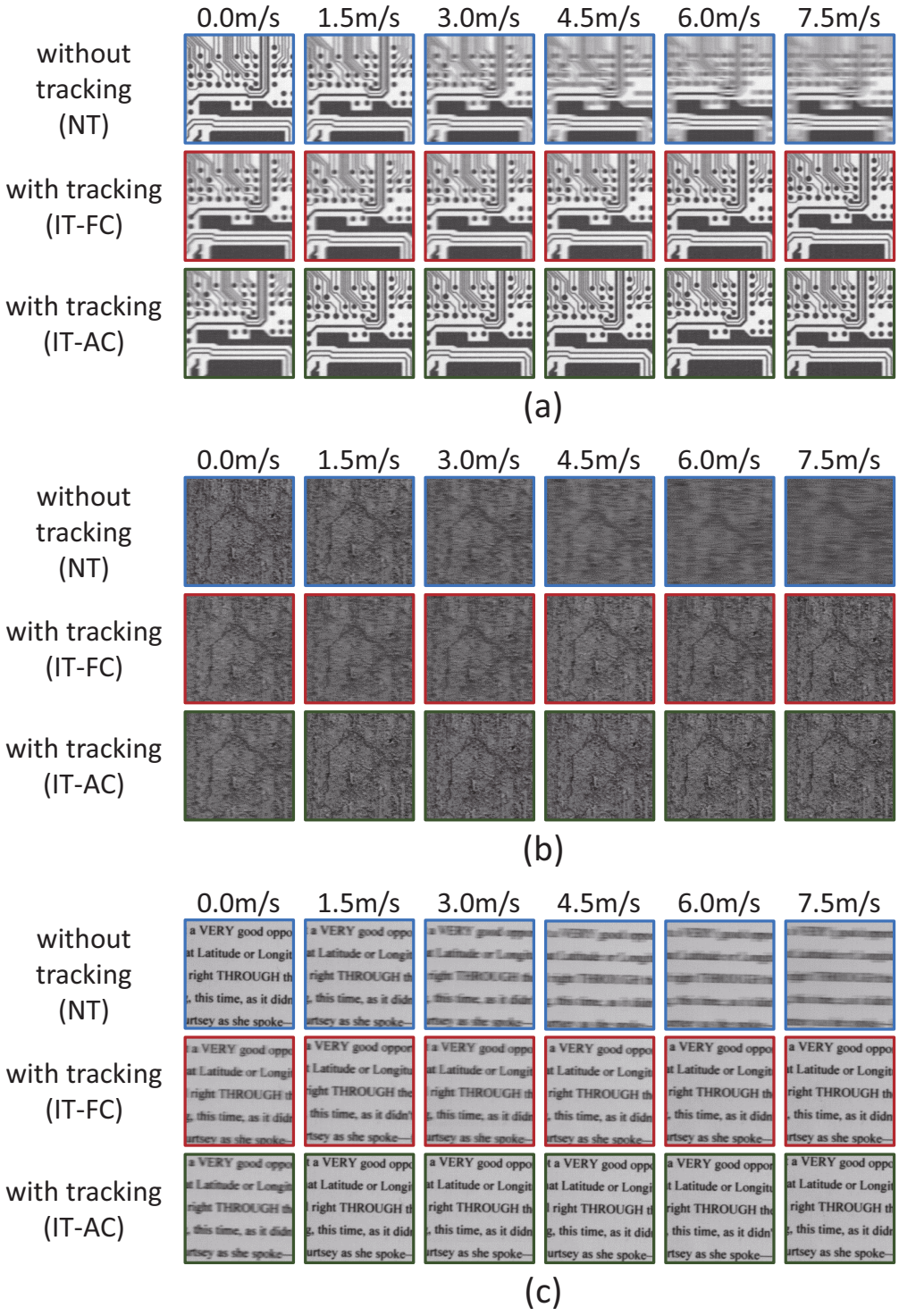


Figure 6.25: Cropped images of complicated patterns in NT, IT-FC and IT-AC at various speeds.

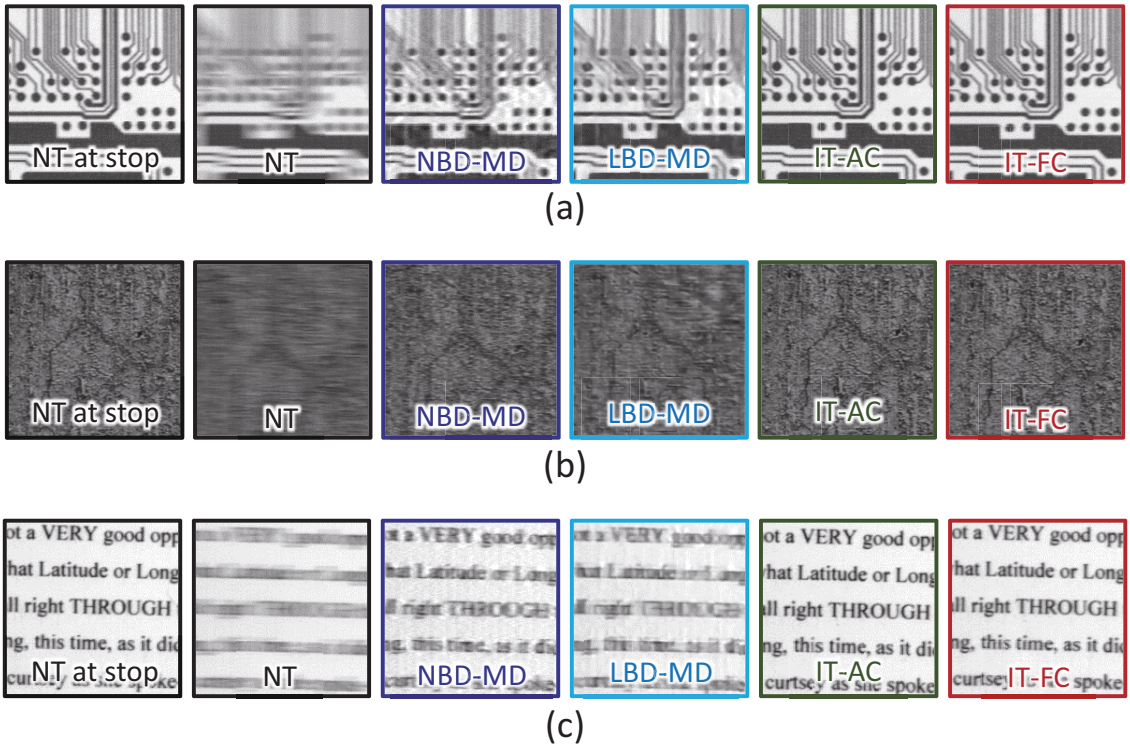


Figure 6.26: Comparison cropped images of several motion-blur-free methods at 7.5 m/s.

width printed on the electronic board pattern, and the crack patterns of less than 0.25-mm width printed on the concrete wall pattern and the text pattern of 2-mm size.

6.4.5 Video Shooting with Frame-capture-event Control from the moving vehicle

Finally, we conducted video capturing experiments by applying the developed system in a vehicle to verify the effectiveness in the situation of capturing the real concrete surface from a running vehicle. In this experiment, the system selected the frame-capture-event control because an unexpected acceleration occurs in these situations.

Figure 6.27 shows an overview of the experimental scene, and a vehicle applied the FPGA-based motion-blur-free HFR video camera. The FPGA-based motion-blur-free HFR video camera system was installed on a seat at the height of 1650 mm with an F-mount 200-mm lens. The shooting distance was set to approximately 4 m between

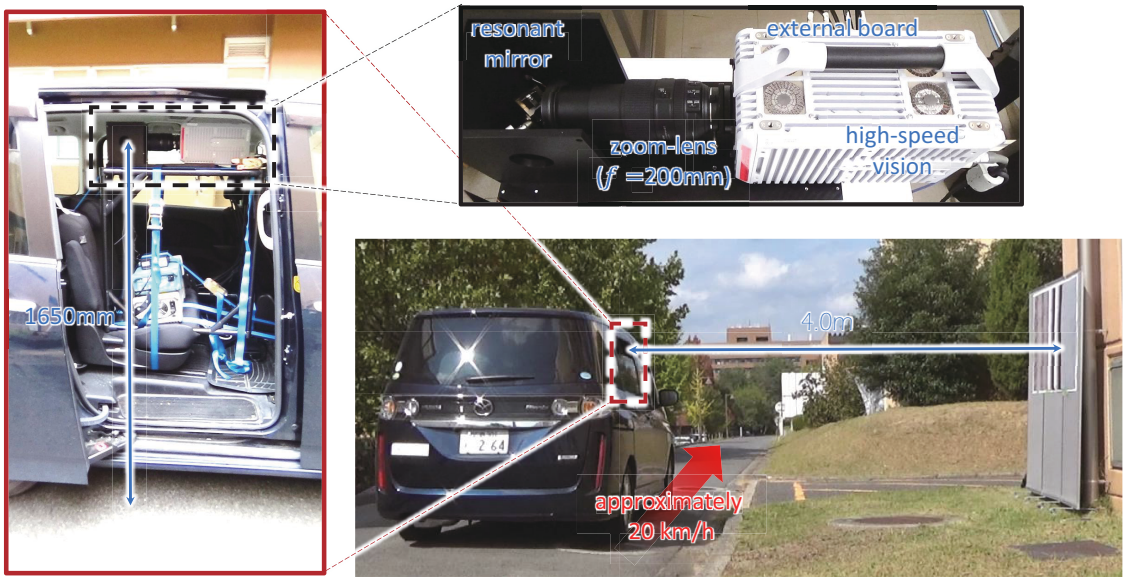


Figure 6.27: Overview of the experimental scene, including the vehicle applied FPGA-based motion-blur-free HFR video camera.

the resonant mirror and the shooting objects. The shooting area was 408×408 mm. The constant mirror amplitude of the resonant mirror was 0.014 deg. The maximum viewpoint velocity of the video camera was 32.5 km/h at the setting camera-to-subject distance. In this experiment, the speed of the running vehicle was set to approximately 20 km/h, whereas the appeared motion-blur at a speed of 20 km/h was around 4 pixels.

Figure 6.28 shows an overview of shooting objects, including (a) test patterns and (b) a concrete wall. The test patterns consist of several actual crack patterns, crack test patterns, and checkered patterns. The images of several actual crack patterns have cracks less than 0.5-mm-width. The crack test patterns have a line with a width of 0.1-mm, 0.2-mm, 0.5-mm, 1.0-mm, and 2.0-mm. The checkered patterns have 4×4 mm squares of alternating black and white were printed.

Figure 6.29(a) shows the result of image sequences by general shooting without tracking (NT), and with tracking by the frame-capture-event control (IT-FC) when it captures the test patterns. Figure 6.29(b) shows the result of captured images when it shoots the actual concrete wall. As seen in the images and edge intensities of the figure, the NT images became blurred in the horizontal direction. In contrast, the frame-capture-

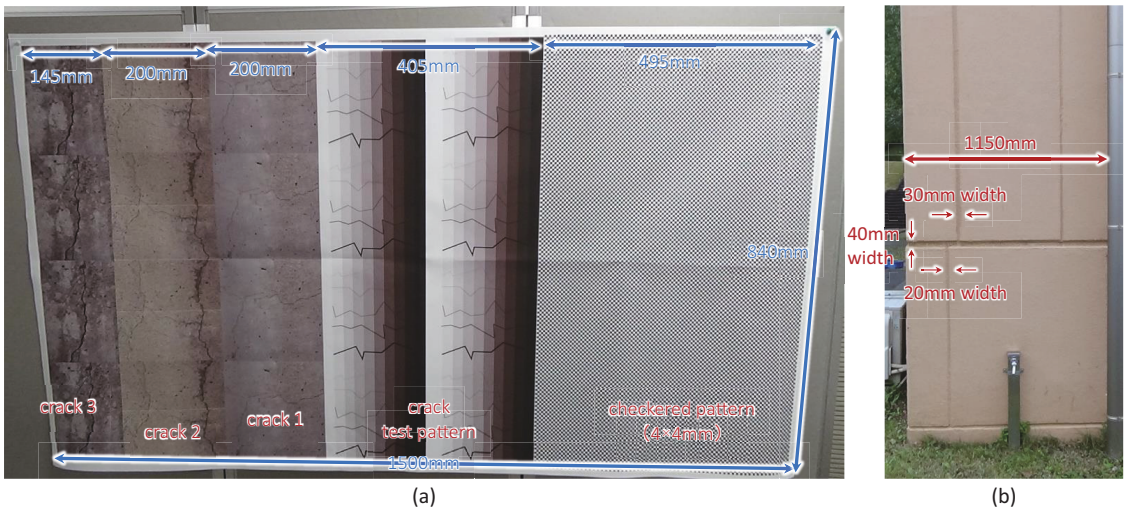


Figure 6.28: Overview of shooting objects; (a) test patterns and (b) a concrete wall.

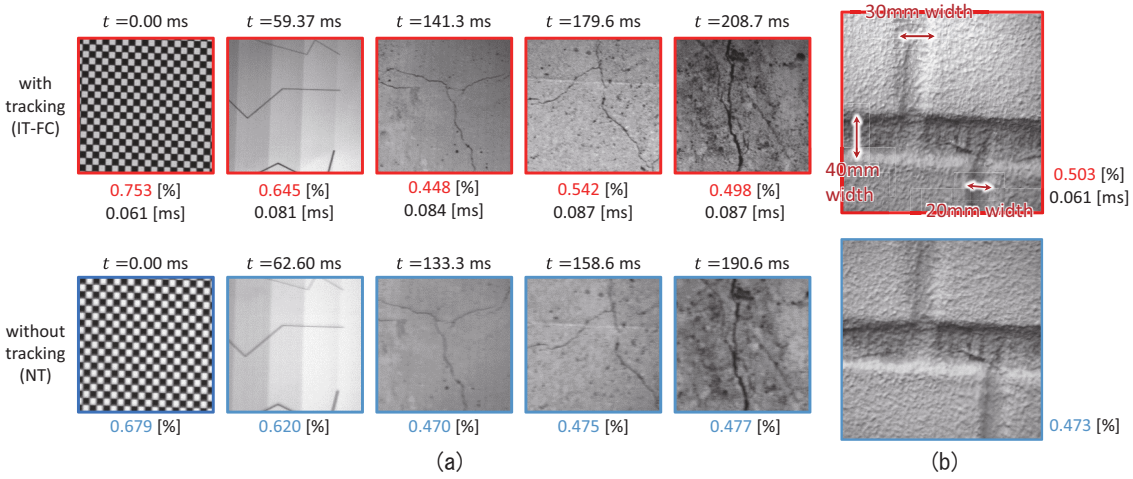


Figure 6.29: (a) 193×193-pixels cropped image sequences when captures test patterns. (b) 430×430-pixels cropped images when captures test patterns.

events τ_d were controlled, and IT-FC images did not become blurred. Figure 6.30 shows the results of normalized edge intensities and frame-capture-event τ_d with tracking when captured the concrete wall. The normalized edge intensities of IT-FC were improved in comparison with NT images. The frame-capture-events τ_d were controlled without relying on the object patterns even when it captured the actual concrete wall surface.

According to these results, our installed motion-blur-free HFR video camera in the vehicle can capture non-blurred images of the actual concrete surface without decreasing

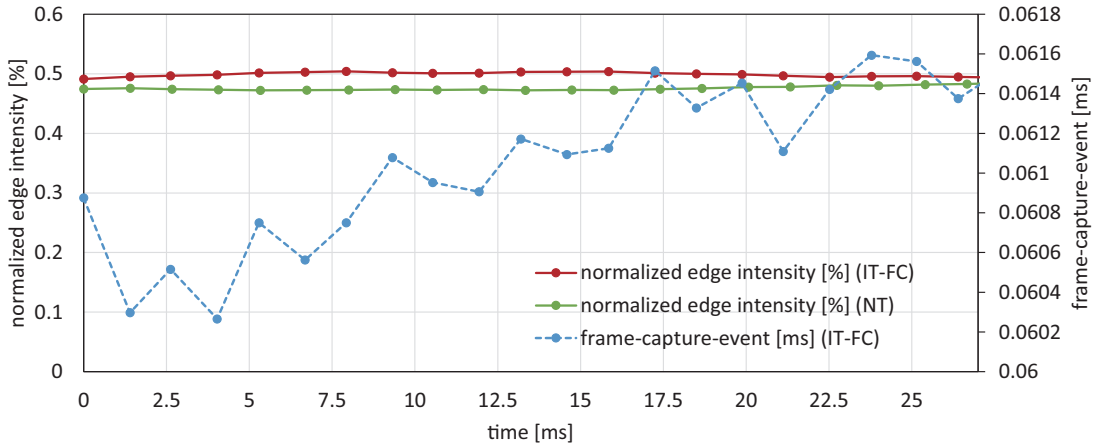


Figure 6.30: Normalized edge intensities and frame-capture-event τ_d with tracking when captured the concrete wall.

the exposure time of the video camera. It indicates our developed vision system can show the effectiveness in the situation of the running vehicle visual inspection, which occurs at the unexpected acceleration.

6.5 Conclusion in this Chapter

In this chapter, we developed the FPGA-based motion-blur-free HFR video camera system and confirmed the effectiveness of the proposed concepts, including the actuator-driven frame-by-frame intermittent tracking and the blur-index-based visual-feedback algorithm. The developed video camera system captures the motion-blur-free HFR video 1024×1024 images of unstable fast-moving scenes at 750 fps without decreasing the exposure time and relying on the object patterns. Several experiments verified the effectiveness of the system, which included the running vehicle experiment for the actual concrete wall. The developed video camera system has two control methods for the viewpoint-velocity during the shutter open; the frame-capture-event control and the amplitude control. The frame-capture-event control realizes the high-frequency feedback at the sub-microsecond level order. In contrast, the percentage of the linear range during the exposure time decreases at lower viewpoint speeds. Its effectiveness described in the acceleration experiments of the conveyor system. The amplitude control realizes non-blur

image shootings compared with the frame-capture-event control at a constant speed, as seen in the several experiments. However, the control responsiveness of the viewpoint velocity is unexpected due to resonance phenomena in the mirror.

On the basis of these results, we plan to extend our motion-blur-free video shooting system to a two-DOF mirror system that can independently control the directions of the pan and tilt mirrors and conduct the verification experiments in where several hundred pixels of motion-blur occur without intermittent mirror tracking. We also plan to introduce an optical device which can control the shooting distance. This is a solution for when the imaging objects move to multiple directions because the actual situation in the fast running car will generate the vibrations.

Chapter 7

Conclusions

In this study, to realize non-blurred image capturing for unstable fast-moving objects without lowering the camera's exposure time, we proposed novel concepts as a motion-blur-free shooting method and developed several motion-blur-free HFR video camera systems consisted of the high-speed vision and the high-speed mirror actuator.

We proposed three concepts for the motion-blur-free high-frame-rate video shooting. The first concept is the camera-driven method with the high-speed mirror actuator for frame-by-frame intermittent tracking. Based on the conventional frame-by-frame intermittent tracking, we extend the method to the mirror-active vision system, which consists of 2-DOF piezo actuators with mirrors. It realizes the motion-blur-free HFR shooting without decreasing the camera's exposure time by alternating the mirror tracking control and back-to-home control according to the camera shutter status. This approach can capture the non-blurred images of the unstable fast-moving object at lower object speeds. The second concept is the actuator-driven method for frame-by-frame intermittent tracking. It realizes the motion-blur-free HFR shooting with long exposure time for fast-moving objects so that the high-speed vision system is controlled in synchronization with the large mirror amplitude of the high-speed mirror actuator. This approach can maximize the mechanical performance of the mirror actuator. The third concept is blur-index-based visual-feedback control. This concept realizes a visual-feedback control for the viewpoint velocity according to the degree of blur in the captured image without relying on the object patterns at frame-by-frame.

The motion-blur-free HFR video camera system, which was implemented the camera-

driven frame-by-frame intermittent tracking method with the high-speed mirror actuator, can capture the 512×512-pixels images at 125 fps without motion-blur. Based on the actuator-driven tracking method and blur-index-based visual-feedback algorithm, the motion-blur-free HFR video camera system was developed by implementing the image processing and the control process on FPGAs. It can capture the non-blurred 1024×1024-pixels images at 750 fps. Its effectiveness was shown through several verification experiments, including the shooting experiment for the table tennis ball and actual concrete walls by installing in the running vehicle. The experimental results show the possibility which can apply our system for solving the motion-blur problems in real fields.

The following issues remain to be solved in the future. In the proposed motion-blur-free HFR video camera system, especially in the actuator-driven approach, the amplitude control with a low-frequency response for viewpoint speed has a limitation. It cannot shoot motion-blur-free videos of an object moving at a rapid time-varying speed in variable directions. In contrast, the frame-capture-event control with a high-responsibility for viewpoint velocity has a limitation as well that the large error occurs at the lower speeds of the shooting objects.

Based on these experimental results and considerations, we plan to combine two viewpoint-velocity control methods to one which can capture non-blurred images with less liner-error during the shutter open and the high responsiveness for the complex patterned objects, as well as to apply the motion-blurring as post-processing to our system to decrease the linear error at the actuator-driven frame-by-frame intermittent tracking using the resonant mirror. We also plan to extend our motion-blur-free video shooting system to a two-DOF mirror system with optical devices that can independently control the pan, tilt, and depth directions by mirrors and lens, and conduct a new practical experiment for much fast-moving scenes such as video shooting investigations from a running vehicle at more than 80 km/h and a flying drone, precise production inspection with a defective product detection on a high-speed factory automation line.

Appendix

I have learned the bottom-up approach for the disadvantaged areas through the TAOYAKA program. This section summarizes the activity of the Onsite-Team-Project in the TAOYAKA program, which was held from October 2017 to July 2018.

As stated above in 1.2, the team in which I joined selected Sera-city in Hiroshima prefecture as the target place at the Onsite-Team-Project. The agriculture of the hilly and mountainous area had an important role in Japan. The hilly and mountainous area occupies 43.3 % of the cultivated land area in Japanese and occupies an important position in the food supply. It plays a significant role in multifaceted functions such as water resource recharge function and flood prevention as well [112]. However, it is facing common issues such as cultivation abandonment and aging of labor in the hilly and mountainous areas in Japan. Sera-city is confronted with similar problems [113]. In this situation, the team started the Onsite-Team-Project in Sera-city and set the project aim that is to support the agriculture of hilly and mountainous areas that have these problems.

In my activity of the Onsite-Team-Project, I proposed a video surveillance system for agricultural supports and surveyed to summarize the system specification. To have a comprehensive understanding of farmers' demands, we investigated two community-based agriculture unions: Kirarikariyama (15 households) and Kurohada (9 households). Through the survey using a questionnaire, I have focused on the rice which is one of the much amount of harvested crops in Japan and the following items: "budget," "frequency to look around," "image quality" and "position in the rice/field to check." In the item of "point in the rice/field to check," the lowest priority is set at point "1", whereas the highest priority is set as point "5".

Figure A.1 shows the questionnaire results in two unions. The frequency to look

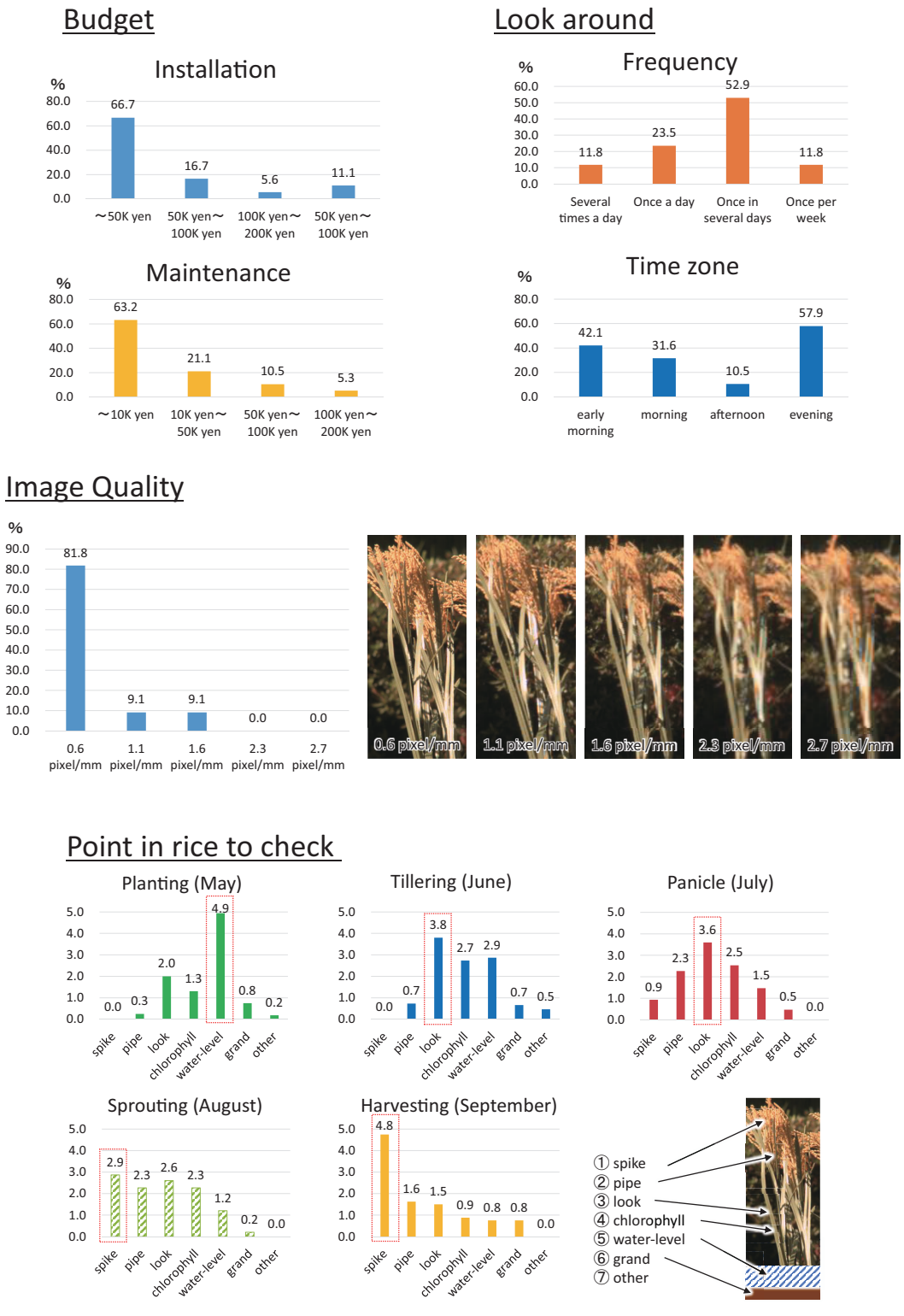


Figure A.1: Questionnaire results of Onsite-Team-Project.

around is once in several days in the evening or early morning. The image quality of cameras is less than averaged 0.74 pixels/mm. The point in rice/field to check is the spike/leaf/water-level, which is monitorable from the sky. Through the result, I have proposed two systems of agriculture support by video surveillance. Figure A.2 shows a summary and an equipment example of my proposal. One of the equipment examples is the Outdoor Security PTZ Camera [114] for the individual; the other is DJI Phantom 4 Pro [115] for community/company. The concept of these systems was selected based on the questionnaire results and specialized for monitoring spike/leaf/water-level from farmer's house.

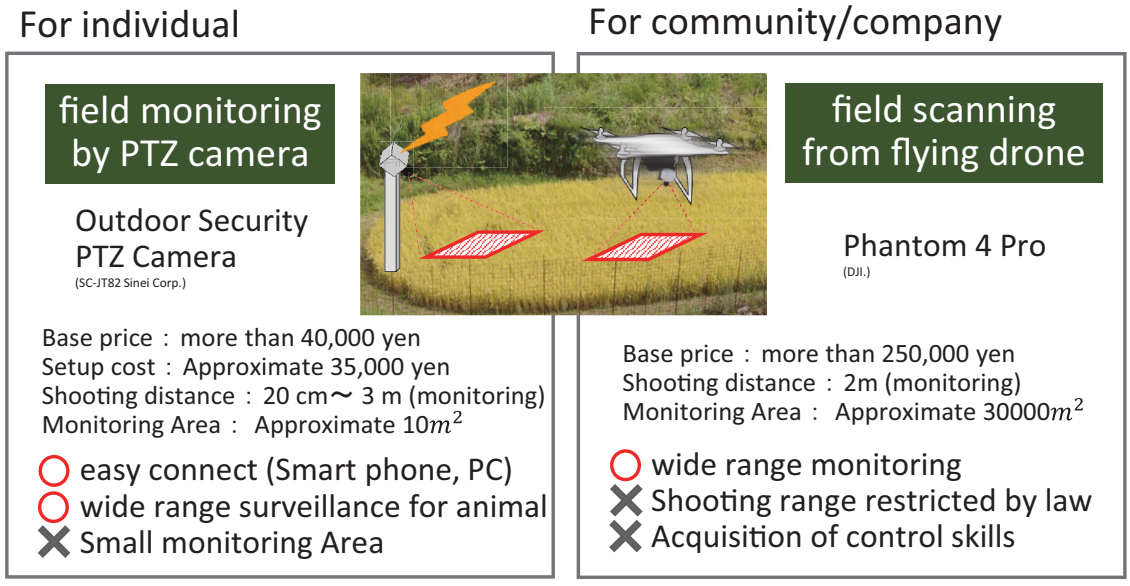


Figure A.2: Proposal of a video surveillance system for agriculture supports.

Bibliography

[1] Ministry of Land, Infrastructure, Transport and Tourism, Japan, https://www.mlit.go.jp/common/001245607.pdf, 2018.

[2] Ministry of Land, Infrastructure, Transport and Tourism, Japan, https://www.mlit.go.jp/common/001228342.pdf, 2018.

[3] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE signal processing magazine*, vol. 13, no. 3, pp. 43–64, 1996.

[4] P. Campisi and K. Egiazarian, "Blind Image Deconvolution: Theory and Applications," CRC Press, 2007.

[5] S. KuKim, J. KiPaik, "Out-of-focus blur estimation and restoration for digital autofocusing system," *Electronics letters*, vol. 34, no. 12, pp. 1217–1219, 1998.

[6] R. Fergus, B. Singh, A. Hertzbann, S.T. Roweis, and W.T. Freeman, "Removing camera shake from a single photograph," *ACM transactions on graphics*, vol. 25, no. 3, pp. 787–794, 2006.

[7] A. Levin, Y. Weiss, F. Durand, and W.T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2657–2664, 2011.

[8] N. Joshi, R. Szeliski, and D.J. Kriegman, "PSF estimation using sharp edge prediction," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, 2008.

[9] S. Cho and S. Lee, "Fast motion deblurring," *ACM transactions on graphics*, vol. 28, no. 5, pp. 145, 2009.

[10] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," *In Proceedings of the European Conference on Computer Vision*, pp. 157–170, 2010.

[11] F. Yang, Y. Huang, Y. Luo, L. Li, and H. Li, "Robust image restoration for motion blur of image sensors," *Sensors*, vol. 16, no. 6, pp. 845, 2016.

[12] N. Xiong, R.W. Liu, M. Liang, D. Wu, Z. Liu, and H. Wu, "Effective alternating direction optimization methods for sparsity-constrained blind image deblurring," *Sensors*, vol. 17, no. 1, pp. 174, 2017.

[13] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 230–240, 2011.

[14] N. Joshi, C.L. Zitnick, R. Szeliski, and D.J. Kriegman, "Image deblurring and denoising using color priors," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1550–1557, 2009.

[15] L. Sun, S. Cho, J. Wang, and J. Hays, "Edge-based blur kernel estimation using patch priors," *In Proceedings of the IEEE Conference on Computational Photography*, pp. 1–8, 2013.

[16] J. Pan, D. Sun, H. Pfister, and M.-H. Yang, "Blind image deblurring using dark channel prior," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1628–1636, 2016.

[17] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM transactions on graphics*, vol. 27, no. 3, pp. 73, 2008.

[18] B. Bascle, A. Blake, and A. Zisserman, "Motion deblurring and super-resolution from an image sequence," *In Proceedings of the European Conference on Computer Vision*, pp. 571–582, 1996.

[19] J. Chen, L. Yuan, C.-K. Tang, L. Quan, "Robust dual motion deblurring," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, 2008.

[20] S. Farsiu, M.D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE transactions on image processing*, vol. 13, no. 10, pp. 1327–1344, 2004.

[21] S.K. Nayar and M. Ben-Ezra, "Motion-based motion deblurring," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 6, pp. 689–698, 2004.

[22] Y.-W. Tai, H. Du, M.S. Brown, and S. Lin, "Image/video deblurring using a hybrid camera," *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8, 2008.

[23] A. Rav-Acha and S. Peleg, "Two motion-blurred images are better than ones," *Pattern recognition letters*, vol. 26, no. 3, pp. 311–317, 2005.

[24] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image deblurring with blurred/noisy image pairs," *ACM transactions on graphics*, vol. 26, no. 3, pp. 1, 2007.

[25] Y.-W. Tai, P. Tan, and M.S. Brown, "Richardson-Lucy deblurring for scenes under a projective motion path," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1603–1618, 2011.

[26] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *International journal of computer vision*, vol. 98, no. 2, pp. 168–186, 2012.

[27] A. Gupta, N. Joshi, C.L. Zitnick, M. Cohen, and B. Curless, "Single image deblurring using motion density functions," *In Proceedings of the European Conference on Computer Vision*, pp. 171–184, 2010.

[28] N. Joshi, S.B. Kang, C.L. Zitnick, and R. Szeliski, "Image deblurring using inertial measurement sensors," *ACM transactions on graphics*, vol. 29, no. 4, pp. 30, 2010.

[29] M.D. Kim and J. Ueda, "Dynamics-based motion de-blurring for a PZT-driven, compliant camera orientation mechanism," *International Journal of Robotics Research*, vol. 34, no. 4–5, pp. 653–673, 2015.

[30] A. Chakrabarti, "A Neural Approach to Blind Motion Deblurring," *European conference on computer vision*, pp. 221–235, 2016.

[31] D. Gong, J. Yang, L. Liu, Y. Zhang, I. Reid, C. Shen, A. V. D. Hengel, and Q. Shi, "From Motion Blur to Motion Flow: A Deep Learning Solution for Removing Heterogeneous Motion Blur," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2319–2328, 2017.

[32] S. Vasu, V.R. Maligireddy, and A.N. Rajagopalan "Non-blind Deblurring: Handling Kernel Uncertainty with CNNs," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3272–3281, 2018.

[33] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, "DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8183–8192, 2018.

[34] H. Kusaka, Y. Tsuchida, and T. Shimohata, "Control technology for optical image stabilization," *SMPTE Motion Imaging Journal*, vol. 111, pp. 609–615, 2002.

[35] B. Cardani, "Optical image stabilization for digital cameras," *IEEE Control Systems Magazine*, vol. 26, pp. 21–22, 2006.

[36] K. Sato, S. Ishizuka, A. Nikami, and M. Sato, "Control techniques for optical image stabilizing system," *IEEE Transactions on Consumer Electronics*, vol. 39, pp. 461–466, 1993.

[37] P. Pournazari, R. Nagamune, and M. Chiao, "A concept of a magnetically-actuated optical image stabilizer for mobile applications," *IEEE Transactions on Consumer Electronics*, vol. 60, pp. 10–17, 2014.

[38] S. Raut, Q. Gu, T. Aoyama, T. Takaki and I. Ishii, "Real-time Optical Image Stabilization Using High Frame Rate Video Sequence at 500 Fps," *Proceedings of the 2015 Conference on Advances In Robotics*, no. 66, pp. 1–6, 2015.

[39] S. Raut, K. Shimasaki, S. Singh, T. Takaki and I. Ishii, "Real-time high-resolution video stabilization using high-frame-rate jitter sensing," *ROBOMECH Journal*, vol. 6, no. 1, pp. 16, 2019.

[40] Q. Hao, X. Cheng, J. Kang, and Y. Jiang, "An image stabilization optical system using deformable freeform mirrors," *Sensors*, vol. 15, no. 1, pp. 1736–1749, 2015.

[41] C.-W. Chiu, P.C.-P. Chao, and D.-Y. Wu, "Optimal design of magnetically actuated optical image stabilizer mechanism for cameras in mobile phones via genetic algorithm," *IEEE Transactions on Magnetics*, vol. 43, no. 6, pp. 2582–2584, 2007.

[42] J.-H. Moon, and S.Y. Jung, "Implementation of an image stabilization system for a small digital camera," *IEEE Transactions on Consumer Electronics*, vol. 54, pp. 206–212, 2008.

[43] M.-G. Song, Y.-J. Hur, N.-C. Park, K.-S. Park, Y.-P. Park, S.-C. Lim, and J.-H. Park, "Design of a voice-coil actuator for optical image stabilization based on genetic algorithm," *IEEE Transactions on Magnetics*, vol. 45, no. 10, pp. 4558–4561, 2009.

[44] M.-G. Song, H.-W. Baek, N.-C. Park, K.-S. Park, T. Yoon, Y.-P. Park, and S.-C. Lim, "Development of small sized actuator with compliant mechanism for optical image stabilization," *IEEE Transactions on Magnetics*, vol. 46, no. 6, pp. 2369–2372, 2010.

[45] T.-H.S. Li, C.-C. Chen, and Y.-T. Su, "Optical image stabilizing system using fuzzy sliding-mode controller for digital cameras," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, pp. 237–245, 2012.

[46] J.H.-S. Wang, K.-F. Qiu, and P.C.-P. Chao, "Control design and digital implementation of a fast 2-degree-of-freedom translational optical image stabilizer for image sensors in mobile camera phones," *Sensors*, vol. 17, no. 10, pp. 2333, 2017.

[47] C.D. Walrath, "Adaptive bearing friction compensation based on recent knowledge of dynamic friction," *Automatica*, vol. 20, no. 6, pp. 717–727, 1984.

[48] B. Ekstrand, "Equations of motion for a two-axes gimbal system," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 3, pp. 1083–1091, 2001.

[49] P.J. Kennedy and R.L. Kennedy, "Direct versus indirect line of sight (LOS) stabilization," *IEEE Transactions on Control Systems Technology*, vol. 11, pp. 3–15, 2003.

[50] X. Zhou, Y. Jia, Q. Zhao, and R. Yu, "Experimental validation of a compound control scheme for a two-axis inertially stabilized platform with multi-sensors in an unmanned helicopter-based airborne power line inspection system," *Sensors*, vol. 16, no. 3, pp. 366, 2016.

[51] Y. Zhang, Y. Xiao, Z. Zhuang, L. Zhou, F. Liu, and Y. He, "Development of a near ground remote sensing system," *Sensors*, vol. 16, no. 5, pp. 648, 2016.

[52] S.-W. Jang, M. Pomplun, G.-Y. Kim, and H.-I. Choi, "Adaptive robust estimation of affine parameters from block motion vectors," *Image and Vision Computing*, vol. 23, no. 14, pp. 1250–1263, 2005.

[53] L. Xu and X. Lin, "Digital image stabilization based on circular block matching," *IEEE Transactions on Consumer Electronics*, vol. 52, no. 2, pp. 566–574, 2006.

[54] W. Chantara, J.-H. Mun, D.-W. Shin, and Y.-S. Ho, "Object tracking using adaptive template matching," *IEIE Trans. Smart Process. Comput.*, pp. 1–9, 2015.

[55] S.-J. Ko, S.-H. Lee, and K.-H. Lee, "Digital image stabilizing algorithms based on bit-plane matching," *IEEE Transactions on Consumer Electronics*, vol. 44, no. 3, pp. 617–622, 1998.

[56] S.-J. Ko, S.-H. Lee, S.-W. Jeon, and E.-S. Kang, "Fast digital image stabilizer based on gray-coded bit-plane matching," *IEEE Transactions on Consumer Electronics*, vol. 45, no. 3, pp. 598–603, 1999.

[57] Y. Shen, P. Guturu, T. Damarla, B.P. Buckles, and K.R. Namuduri, "Video stabilization using principal component analysis and scale invariant feature transform in particle filter framework," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 3, pp. 1714–1721, 2009.

[58] J. Xu, H. Chang, S. Yang, and M. Wang, "Fast feature-based video stabilization without accumulative global motion estimation," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 993–999, 2012.

[59] S. Liu, L. Yuan, P. Tan, and J. Sun, "Bundled camera paths for video stabilization," *ACM transactions on graphics*, vol. 32, no. 4, pp. 78, 2013.

[60] S.-K. Kim, S.-J. Kang, T.-S. Wang, and S.-J. Ko, "Feature point classification based global motion estimation for video stabilization," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 1, pp. 267–272, 2013.

[61] X. Cheng, Q. Hao, and M. Xie, "A comprehensive motion estimation technique for the improvement of EIS Methods based on the SURF algorithm and Kalman filter," *Sensors*, vol. 16, no. 4, pp. 486, 2016.

[62] S. Jeon, I. Yoon, J. Jang, S. Yang, J. Kim, and J. Paik, "Robust video stabilization using particle keypoint update and l_1-optimized camera path," *Sensors*, vol. 17, no. 2, pp. 337, 2017.

[63] J.-Y. Chang, W.-F. Hu, M.-H. Cheng, and B.-S. Chang, "Digital image translational and rotational motion stabilization using optical flow technique," *IEEE Transactions on Consumer Electronics*, vol. 48, no. 1, pp. 108–115, 2002.

[64] Y. Matsushita, E. Ofek, W. Ge, X. Tang, and H.-Y. Shum, "Full-frame video stabilization with motion inpainting," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 7, pp. 1150–1163, 2006.

[65] W. Xu, X. Lai, D. Xu, and N.A. Tsoligkas, "An integrated new scheme for digital video stabilization," *Advances in Multimedia*, 2013.

[66] S. Pathak, A. Moro, H. Fujii, A. Yamashita, and H. Asama, "Spherical video stabilization by estimating rotation from dense optical flow fields," *Journal of Robotics and Mechatronics*, vol. 29, no. 3, pp. 566–579, 2017.

[67] H.E. Edgerton and K.J. Germeshausen, "Stroboscopic-light high-speed motion pictures," *Journal of the Society of Motion Picture Engineers*, vol. 23, no. 5, pp. 284–298, 1934.

[68] D. Bradley, B. Atcheson, I. Ihrke, and W. Heidrich, "Synchronization and rolling shutter compensation for consumer video camera arrays," *In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1–8, 2009.

[69] F. Boden, K. Bodensiek, and B. Stasicki, "Application of image pattern correlation for non-intrusive deformation measurements of fast rotating objects on aircrafts," *In Proceedings of the Fourth International Conference on Experimental Mechanics*, pp. 75222S, 2009.

[70] C. Theobalt, I. Albrecht, J. Haber, M. Magnor, and H.-P. Seidel, "Pitching a baseball: Tracking high-speed motion with multi-exposure images," *ACM transactions on graphics*, vol. 23, no. 3, pp. 540–547, 2004.

[71] F.H. Borsato, F.O. Aluani, and C.H. Morimoto, "A fast and accurate eye tracker using stroboscopic differential lighting," *In Proceedings of the IEEE International Conference on Computer Vision Workshop*, pp. 110–118, 2015.

[72] Y. Watanabe, T. Komura, and M. Ishikawa, "955-fps real-time shape measurement of a moving/deforming object using high-speed vision for numerous-point analysis," *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pp. 3192–3197, 2007.

[73] I. Ishii, T. Taniguchi, R. Sukenobe, and K. Yamamoto, "Development of high-speed and real-time vision platform, H3 Vision," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3671–3678, 2009.

[74] I. Ishii, T. Tatebe, Q. Gu, Y. Moriue, T. Takaki, and K. Tajima, "2000 fps Real-time vision system with high-frame-rate video recording," *In Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1536–1541, 2010.

[75] T. Yamazaki, H. Katayama, S. Uehara, A. Nose, M. Kobayashi, S. Shida, M. Odahara, K. Takamiya, Y. Hisamatsu, S. Matsumoto, et al. "A 1ms high-Speed vision chip with 3D-stacked 140GOPS column-parallel PEs for spatio-temporal image processing," *In Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 82–83, 2017.

[76] I. Ishii, T. Taniguchi, K. Yamamoto, and T. Takaki, "High-frame-rate optical flow system," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 1, pp. 105–112, 2011.

[77] I. Ishii, T. Tatebe, Q. Gu, and T. Takaki, "Color-histogram-based tracking at 2000 fps," *Journal of Electronic Imaging*, vol. 21, no. 1, pp. 013010, 2012.

[78] Q. Gu, T. Takaki, and I. Ishii, "Fast FPGA-based multiobject feature extraction," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 1, pp. 30–45, 2013.

[79] Q. Gu, S. Raut, K. Okumura, T. Aoyama, T. Takaki, and I. Ishii, "Real-time image mosaicing system using a high-frame-rate video sequence," *Journal of Robotics and Mechatronics*, vol. 27, no. 1, pp. 12–23, 2015.

[80] I. Ishii, T. Ichida, Q. Gu, and T. Takaki, "500-fps face tracking system," *Journal of real-time image processing*, vol. 8, no. 14, pp. 379–388, 2013.

[81] A. Namiki, K. Hashimoto, and M. Ishikawa, "A hierarchical control architecture for high-speed visual servoing," *International Journal of Robotics Research*, vol. 22, no. 10–11, pp. 873–888, 2003.

[82] T. Senoo, A. Namiki, and M. Ishikawa, "Ball control in high-speed batting motion using hybrid trajectory generator," *In Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1762–1767, 2006.

[83] A. Namiki and N. Ito, "Ball catching in Kendama game by estimating grasp conditions based on a high-speed vision system and tactile sensors," *In Proceedings of the IEEE Conference on Humanoid Robots*, pp. 634–639, 2014.

[84] T. Aoyama, T. Takaki, T. Miura, Q. Gu, and I. Ishii, "Realization of flower stick rotation using robotic arm," *In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5648–5653, 2015.

[85] M. Jiang, T. Aoyama, T. Takaki, and I. Ishii, "Pixel-level and robust vibration source sensing in high-frame-rate video analysis," *Sensors*, vol. 16, no. 11, pp. 1842, 2016.

[86] M. Jiang, Q. Gu, T. Aoyama, T. Takaki, and I. Ishii, "Real-time vibration source tracking using high-speed vision," *IEEE Sensors Journal*, vol. 17, no. 5, pp. 1513–1527, 2017.

[87] K. Shimasaki, M. Jiang, T. Takaki, I. Ishii, and K. Yamamoto, "HFR-Video-Based Honeybee Activity Sensing Using Pixel-Level Short-Time Fourier Transform," *IEEE Sensors*, pp. 1–4, 2018.

[88] H. Oku, I. Ishii, and M. Ishikawa, "Tracking a protozoon using high-speed visual feedback," *In Proceedings of the International IEEE-EMBS Conference on Microtechnologies in Medicine and Biology*, pp. 156–159, 2000.

[89] S. Sakuma, K. Kuroda, C.-H.D. Tsai, W. Fukui, F. Arai, and M. Kaneko, "Red blood cell fatigue evaluation based on the close-encountering point between extensibility and recoverability," *Lab on a Chip*, vol. 14, no. 6, pp. 1135–1141, 2014.

[90] Q. Gu, T. Aoyama, T. Takaki, and I. Ishii, "Simultaneous vision-based shape and motion analysis of cells fast-flowing in a microchannel," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 1, pp. 204–215, 2015.

[91] Q. Gu, T. Kawahara, T. Aoyama, T. Takaki, I. Ishii, A. Takemoto, and N. Sakamoto, "LOC-based high-throughput cell morphology analysis system," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 4, pp. 1346–1356, 2015.

[92] H. Yang, Q. Gu, T. Aoyama, T. Takaki, and I. Ishii, "Dynamics-based stereo visual inspection using multidimensional modal analysis," *IEEE Sensors Journal*, vol. 13, no. 12, pp. 4831–4843, 2013.

[93] K. Okumura, K. Yokoyama, H. Oku, and M. Ishikawa, "1 ms auto pan-tilt—Video shooting technology for objects in motion based on Saccade Mirror with background subtraction," *Advanced Robotics*, vol. 29, no. 7, pp. 457–468, 2015.

[94] L. Li, T. Aoyama, T. Takaki, I. Ishii, H. Yang, C. Umemoto, H. Matsuda, M. Chikaraishi, and A. Fujiwara, "Vibration distribution measurement using a high-speed multithread active vision," *In Proceedings of the IEEE Conference on Advanced Intelligent Mechatronics (AIM)*, pp. 400–405, 2017.

[95] H.-S. Wong, Y.L. Yao, and E.S. Schlig, "TDI charge-coupled devices: Design and applications," *IBM Journal of research and development*, vol. 36, no. 1, pp. 83–106, 1992.

[96] G. Lepage, J. Bogaerts, and G. Meynants, "Time-Delay-Integration Architectures in CMOS Image Sensors," *IEEE Transactions on Electron Devices*, vol. 56, no. 11, pp. 2524–2533, 2009.

[97] C. Yin, T. Liao, K.L. Liu, C.C. Kao, C.F. Chiu, and C.C. Hsieh, "A 32-Stage 15-b Digital Time-Delay Integration Linear CMOS Image Sensor With Data Prediction Switching Technique," *IEEE Transactions on Electron Devices*, vol. 64, no. 3, pp. 1167–1173, 2017.

[98] S.I. Raiter, A. Stern, O. Hadar, and N.S. Kopeika, "Image restoration from camera vibration and object motion blur in infrared staggered time-delay and integration systems," *Optical Engineering*, vol. 42, no. 11, pp. 3253–3265, 2003.

[99] J. Wu, Z. Zheng, H. Feng, Z. Xu, and Y. Chen, "Restoration of TDI camera images with motion distortion and blur," *Optics and Laser Technology*, vol. 42, no. 8, pp. 1198–1203, 2010.

[100] T. Ueno, Q. Gu, T. Aoyama, T. Takaki, I. Ishii, and T. Kawahara, "Motion-blur-free microscopic video shooting based on frame-by-frame intermittent tracking," *In Proceedings of the IEEE Conference on Automation Science and Engineering*, pp. 837–842, 2015.

[101] M. Inoue, Q. Gu, T. Aoyama, T. Takaki, and I. Ishii, "An intermittent frame-by-frame tracking camera for motion-blur-free video shooting," *In Proceedings of the 2015 IEEE/SICE International Symposium on System Integration*, pp. 241–246, 2015.

[102] M. Inoue, M. Jiang, Y. Matsumoto, T. Takaki, and I. Ishii, "Motion-blur-free video shooting system based on frame-by-frame intermittent tracking," *ROBOMECH Journal*, vol. 4, no. 1, pp. 28, 2017.

[103] T. Hayakawa, T. Watanabe, and M. Ishikawa, "Real-time high-speed motion blur compensation system based on back-and-forth motion control of galvanometer mirror," *Opt. Express*, vol. 23, no. 25, pp. 31648–31661, 2015.

[104] T. Hayakawa and M. Ishikawa, "Development of motion-blur-compensated high-speed moving visual inspection vehicle for tunnels," *Int. J. Civ. Struct. Eng. Res.*, vol. 5, pp. 151–155, 2015.

[105] A. Sharma, K. Shimasaki, Q. Gu, J. Chen, T. Aoyama, T. Takaki, I. Ishii, K. Tamura, and K. Tajima, "Super high-speed vision platform for processing 1024×1024 images in real time at 12500 fps," *IEEE/SICE International Symposium on System Integration (SII)*, pp. 544–549, 2016.

[106] M. Inoue, Q. Gu, M. Jiang, T. Takaki, I. Ishii, and K. Tajima, "Motion-blur-free High-speed Video Shooting with Frame-timing Control," *12th International Conference on Sensing Technology (ICST)*, pp. 172-177, 2018.

[107] J. G. Ziegler and N. B. Nichols, "Optimum Settings for Automatic Controllers," *Trans. ASME*, vol. 64, no. 11, 1942.

[108] S. Tavakoli and M. Tavakoli, "Optimal tuning of PID controllers for first order plus time delay models using dimensional analysis," *4th International Conference on Control and Automation Proceedings*, pp. 942–946, 2003.

[109] S. E. Wright, A. W. Mahoney, K. M. Popek and J. J. Abbott, "The spherical-actuator-magnet manipulator: A permanent-magnet robotic end-effector," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1013–1024, 2017.

[110] A. Levin, P. Sand, T. S. Cho, F. Durand, and W. T. Freeman, "Motion-invariant photography," *ACM Trans. Graph.*, vol. 27, no. 71, pp. 1–9, 2008.

[111] X. Tao, H. Gao, X. Shen, J. Wang, Jue and J. Jia, "Scale-recurrent Network for Deep Image Deblurring," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8174–8182, 2018.

[112] Rural Promotion Bureau in Ministry of Agriculture, Forestry and Fisheries of Japan, "Situation of regional agriculture in inter-mountainous areas," http://www.maff.go.jp/j/study/other/cyusan_taisaku/32/pdf/data1.pdf, 2009.

[113] Sera Town Office, "Population vision," http://www.town.sera.hiroshima.jp/data/open/cnt/3/952/1/matihitoshigotozinkou.pdf, 2015.

[114] SecuSTATION, "SC-JT support," https://secu.jp/support/jt.html, 2019.

[115] DJI, "Phantom 4 Pro," https://www.dji.com/jp/phantom-4-pro, 2019.

Acknowledgment

Firstly, I wish to appreciate my advisor, **Prof. Idaku Ishii**, who provided me an opportunity to join our laboratory and led me into academic research during the six years. He taught me a lot of skills and attitudes for academic research in aspects of a new research consideration, a research process method, a research summarization such as presentations and papers. I want to express my gratitude again.

Besides my advisor, I would like to express my gratitude to **Dr. Takeshi Takaki**, **Dr. Mingjun Jiang**, **Dr. Qingyi Gu**, **Dr. Shaopeng Hu**, **Mr. Sushil Raut**, and **Mr. Kohei Shimasaki** Their invaluable suggestions helped me overcome the unfamiliarity with a new experimental environment when I initially joined our laboratory. I would also like to express my heartfelt gratitude to **Ms. Yukari Kaneyuki** (educational administrator), **Ms. Rumi Horiuchi**, **Ms. Michiko Kanzaki**, and **Ms. Arisa Tomura** (laboratory secretary). They were my most reliable staff in our institution; I received thoughtful attention both in my study and life from them. I also want to express my sincere thanks to the bachelor, master, and doctoral students in Robotics Laboratory for their help in life and my research. I want to thank the teachers, clerical staff, and doctoral students in the "TAOYAKA Program" who had the opportunity to experience many fieldwork and events that other doctoral students do not have as well.

Finally, I want to express my profound gratitude to my family for their moral support and warm encouragement for more than the past two decades throughout my life in general.

March, 2020

Michiaki Inoue