

Distributed Optimization Framework for Shadow Removal in Multi-Projection Systems

J. Tsukamoto¹, D. Iwai² and K. Kashima¹

¹Kyoto University, Japan

²Osaka University, Japan

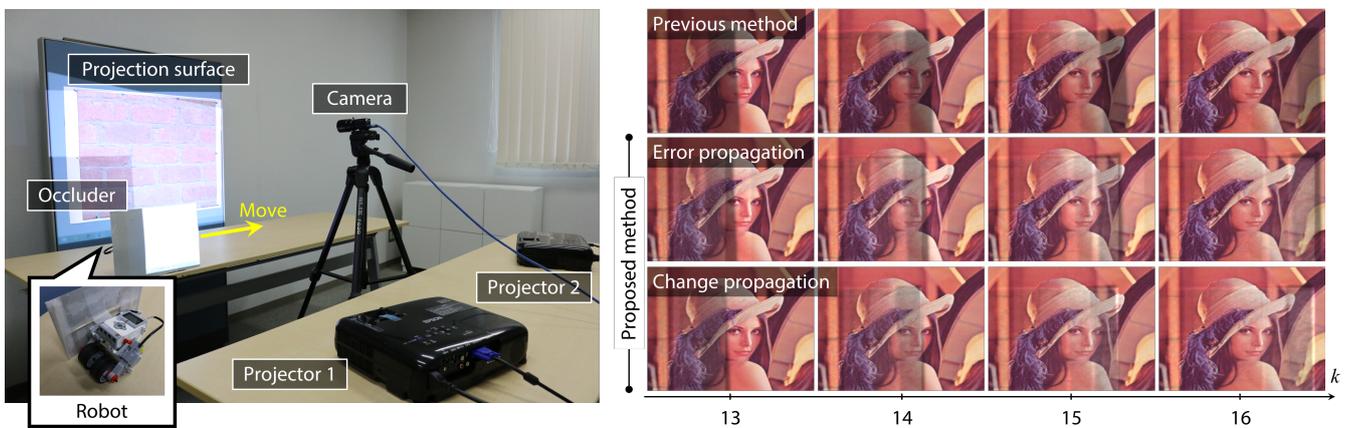


Figure 1: Shadow removal in a distributed projector-camera system: (left) experimental system where two overlapping projectors were applied for removing shadows caused by a moving occluder; (right) comparison of projected results in the lenna condition (see Sec. 5.1): (top row) the results by the previous method [TIK15], (middle row) the results by the error propagation method, (bottom row) the results by the change propagation method.

Abstract

This paper proposes a novel shadow removal technique for cooperative projection system based-on spatio-temporal prediction. In our previous work, we proposed a distributed feedback algorithm, which is implementable in cooperative projection environments subject to data transfer constraints between components. A weakness of this scheme is that the compensation is conducted in each pixel independently. As a result, spatio-temporal information of the environmental change cannot be utilized even if it is available. In view of this, we specifically investigate the situation where some of projectors are occluded by a moving object whose one-frame-ahead behavior is predictable. In order to remove the resulting shadow, we propose a novel error propagating scheme that is still implementable in a distributed manner, and enables us to incorporate the prediction information of the obstacle. It is demonstrated theoretically and experimentally that the proposed method significantly improves the shadow removal performance comparison to the previous work.

Categories and Subject Descriptors (according to ACM CCS): I.4.0 [Image Processing and Computer Vision]: General—Image displays

1. Introduction

Shadow removal is an important fundamental technology for interactive front projection systems such as spatial augmented real-

ity (SAR), in which shadows of projected images are frequently casted, because users dynamically move between projectors and surfaces, and consequently, block the projected images easily and

unexpectedly. These shadows must be removed, as they potentially occlude important projected information and detract from the users' visually immersive experiences. One of the most promising shadow removal solutions is to employ multiple overlapping projectors, in which at least a projector visible from a shadow area projects images onto it instead of the other occluded projectors. Recently, ultra-short throw projectors are commercially available, by which no shadow occurs for a flat projection surface, because the distance between the projector and screen is very short. On the other hand, the goal of this paper is to remove shadows in general SAR applications where not only a single flat surface but also multiple non-planar surfaces located at different depths are assumed to be projection surfaces [BR05]. Shadows can be caused in these situations even when we use ultra-short throw projectors.

A multi-projection system should be scalable so that it accommodates the demand of increasing computational costs and communications traffic when the number of projector nodes increases. In other words, it is important to have a Plug-and-Play capability by which a newly added (plug-in) projector node is automatically connected to the system's network. The system also should be robust for the failure (plug-out) of projector nodes at runtime. However, these issues have not been carefully considered in previous works of the shadow removal researches, which consequently have applied a centralized control approach in which a host node computes projection images for all projector nodes. To achieve high scalability, robustness and Plug-and-Play capability as mentioned above, we previously proposed to apply a distributed optimization algorithm based on distributed control theory to radiometric compensation for a cooperative overlapping projection system [TIK15]. Although our previous framework worked for shadow removal, it is not optimized for a dynamic occluder, which is more common in interactive systems than a static one.

In this paper, we propose a novel shadow removal technique for our multiple overlapping projection system by extending the previously proposed distributed optimization framework. The centre of the contributions of the paper is to apply an error propagation scheme that propagates the current error information (which is mainly caused by a shadow) to the next frame according to the motion of the shadow. The new technique improves the tracking responsiveness (i.e., speed of convergence) to realize faster removal of dynamic shadows than the previous method. Based on a control theoretic approach, we evaluate the improvement through an explicit analysis of the linearized error dynamics. We also conduct simulation and physical projection experiments to validate the shadow removal performance in an environment where a shadow area moves according to the occluder's locomotion.

To summarize, this paper makes the following contributions:

- We extend the previously proposed distributed optimization-based radiometric compensation technique to significantly improve the shadow removal performance by applying an error propagation scheme.
- We provide the theoretical performance limit of the proposed technique in terms of the tracking responsiveness using a control theoretic approach, and show how it outperforms the previous technique.
- Through simulation and physical projection experiments, we

demonstrate the feasibility of the proposed method in a dynamic shadow environment, in terms of the removal accuracy and the speed of convergence.

2. Related Work

Following the pioneering work by Sukthankar et al. [SCS01], several approaches have been proposed for shadow removal in multi-projection systems. All these methods employ multiple overlapping projectors, and single or multiple cameras to capture either a shadow on the projection surface, or an occluder. Sukthankar et al. proposed to remove shadows from a projected result assuming a static image is displayed [SCS01]. Their technique applied a feedback process where, at each iteration, it generates a projection image so that the error of the current projected result from the target appearance is minimized.

As another approach, Jaynes et al. proposed to apply shadow area detection technique to compute projection images that can remove shadow areas at each frame without relying on a feedback method [JWS04, JWS*01]. Their methods compare a captured scene with a predicted projection result to find a shadow region that was then illuminated by unoccluded projectors. Shadow area detection approach was also proposed by Sugaya et al., who assigned different intensity values to each projector to identify occluded projectors from a single captured image of shadows [SMK10]. In another work, synthetic aperture technique using multiple cameras was used to detect shadow areas, which work even when a part of cameras cannot observe a projection surface due to occluders [INS14].

Instead of shadow area detection, occluder detection has been also applied by some researchers. Summet et al. detected an occluder by illuminating the scene with infrared (IR) lights and capturing the projection surface with an IR camera [SFC*07]. Audet et al. proposed recovering the depth information of an occluder with two cameras [AC07]. Once the depth information was obtained, the shadow area of the occluder on the projection surface could be estimated from the geometric relationships among the projectors, surface, and occluder.

Although these methods worked well, they heavily rely on a host computer that has centralized control over all projectors. Therefore, they are not suitable for our goal that is to realize a highly scalable, robust, and Plug-and-Play capable shadow removal, as mentioned in the first section. In our previous work [TIK15], we built a radiometric compensation technique that meets our requirements using a distributed optimization algorithm [NO09]. Similar approach was also proposed by other researchers [AK10, ASUK14]. Although our previous technique [TIK15] was not explicitly designed to remove shadows as other previous radiometric compensation techniques [BIWG08], it can collaterally remove shadows as a consequence of the radiometric compensation of multiple overlapping projectors. However, it works well only when the shadow areas do not move.

In this paper, we extend our previous distributed optimization framework so that the moving shadow areas are removed more effectively. The proposed technique in this paper assumes a distributed cooperative system consisting of a single camera node and

$i \in \mathcal{L} := \{1, 2, \dots, n\}$	projector index
$j \in \mathcal{J} := \{1, 2, \dots, m\}$	pixel index
$k \in \mathbb{N}$	frame index
$x_i^j[k] \in [0, 1]$	input image
$y^j[k] \in \mathbb{R} (j \in \mathcal{J})$	projected result image
$p_i^j \in \mathbb{R}_+ (i \in \mathcal{L}, j \in \mathcal{J})$	form factor
$r^j \in \mathbb{R}_+ (j \in \mathcal{J})$	target image
$d^j[k] \in \mathbb{R} (j \in \mathcal{J})$	environmental light

Figure 2: Notation.

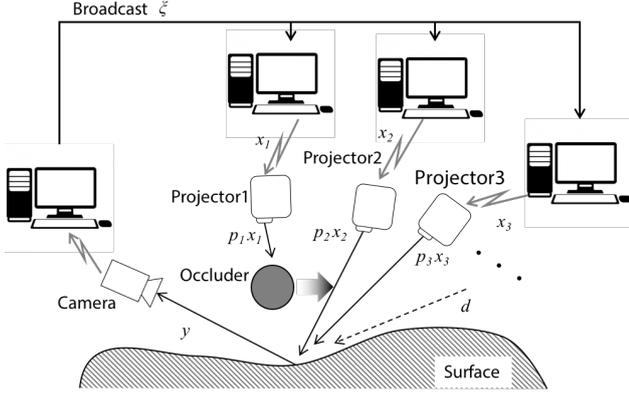


Figure 3: System architecture.

multiple projector nodes. It shows a clear contrast to the existing shadow removal results that the following information limitation is taken into an explicit account: (1) each projector node does not have any information about other projector nodes, shadow areas, and target images, (2) the camera node does not have any information about the projector nodes either, while having the target images, and (3) only a broadcast communication from the camera node to the projector nodes is allowed to suppress the data transfer bandwidth.

3. Problem Formulation

This section describes the problem formulation of this research, which is based on the previously proposed distributed feedback algorithm of radiometric compensation [TIK15]. We apply the Bimber's simple linear model [BEK05] in this paper, while other more complex linear models can be also applied (e.g., [GPNB04]). As with most of radiometric compensation methods, we assume our projection objects as arbitrarily shaped and textured but limited to Lambertian surfaces.

3.1. System architecture and problem formulation

We consider a projector-camera system composed of n projector nodes and one camera node that is regarded as an eye of a human observer. Projected images from the projector nodes are overlapped each other on a projection surface. At a time or frame k , each projector node i has an input image $x_i^j[k] \in [0, 1]$ and form

factor p_i^j for the j -th pixel. In this setting, according to the linear model [BEK05], the camera observation at the j -th pixel can be represented as

$$y^j[k] = \sum_i p_i^j x_i^j[k] + d^j[k], \quad (1)$$

where $d^j[k]$ is environmental light (e.g. ambient light and black offset) (Fig. 3). The form factor p_i^j represents an attenuation ratio from the input pixel intensity to the projected result. The form factor is affected by the reflectance of the surface, the distance between the projector to the surface, incident angle of projected light, and so on. The calibration of the form factor is performed normally by projecting a uniform white image and capturing the reflection on a projection surface, which should be done only once when each projector is newly added to the system.

Suppose the camera node has a target image r^j at each frame k . In order to project the target image accurately, we minimize the following objective function

$$G(x[k]) = \frac{1}{2} \sum_j (e^j[k])^2, \quad x[k] = (x_i^j[k]), \quad i \in \mathcal{L}, \quad j \in \mathcal{J}, \quad (2)$$

which is the sum of the squared error

$$e^j[k] = r^j - y^j[k] = r^j - \sum_{i=1}^n p_i^j x_i^j[k] - d^j[k], \quad (3)$$

Symbols of this model are listed in Fig. 2.

3.2. Distributed feedback algorithm of radiometric compensation

To achieve high scalability and robustness, we assume that $x_i^j[k]$ and p_i^j are available only for the projector node i , and that r^j and $y^j[k]$ are available only for the camera node. Thanks to the specific structure of Eq. (2), we can optimize $x_i^j[k]$ in such a distributed setting. The gradient of Eq. (2) is given by

$$\frac{\partial G}{\partial x_i^j[k]} = -p_i^j e^j[k]. \quad (4)$$

Therefore, if $e^j[k]$ is available, each projector node can compute the gradient without any other information about the other projectors. For implementation, the camera node can compute and broadcast

$$\xi^j[k] = K e^j[k] = K(r^j - y^j[k]) \quad (5)$$

where K is a positive constant. Then, each projector updates $x_i^j[k]$ by

$$x_i^j[k+1] = \mathcal{P}[x_i^j[k] + p_i^j \xi^j[k]] \quad (6)$$

$$\mathcal{P}[\bar{x}] := \begin{cases} \bar{x}, & 0 \leq \bar{x} \leq 1, \\ 1, & \bar{x} > 1, \\ 0, & \bar{x} < 0, \end{cases} \quad \bar{x} \in \mathbb{R}, \quad (7)$$

where \mathcal{P} represents a clipping (or projection) process to avoid saturated pixel values. This algorithm is equivalent to a projected gradient method to minimize G where K works as a stepsize. Actually, it can be proven that this update rule with sufficiently small K makes $x_i^j[k]$ converge to the optimal value in a suitable sense [Ste15].

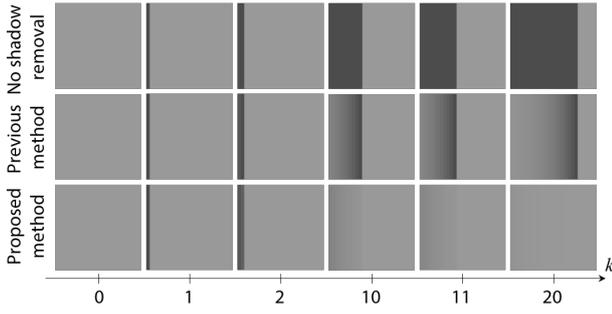


Figure 4: Simulated shadow removal results in a simple case.

To analyze the dynamical stability, let us consider the following difference equation where the saturation \mathcal{P} and disturbance d is ignored:

$$e^j[k+1] = r^j - y^j[k+1] \quad (8)$$

$$= (1 - \kappa^j K) e^j[k] \quad (9)$$

with

$$\kappa^j := \sum_i (p_i^j)^2. \quad (10)$$

Therefore, a necessary and sufficient condition for this system to be stable is given by

$$0 < K < \frac{2}{\kappa^j}. \quad (11)$$

As described above, once $\xi^j[k]$ is broadcasted to all the projector nodes from the camera node, the projector nodes can compute input images without other communications among the nodes in the system. This is a completely distributed process, and furthermore, it is theoretically guaranteed that removal or addition (i.e., plug-out or plug-in) of projector nodes at run time does not affect the convergence performance. Please refer to [TIK15] for more detailed information about the algorithm.

3.3. Problem setting

Now, we explain the situation of our interest in this paper. Suppose that an object occludes the projection of some projectors, but not the camera observation; see Fig. 3. We assume that shadow regions in the captured image coordinate are predictable one frame ahead, e.g., by using optical flows [Far03]. Our goal is to remove shadow caused by the occlusion.

Let us see what happens if the aforementioned distributed feedback algorithm (see Sec. 3.2) is applied to this situation. Without loss of generality, we consider a simple example case as follows. The target appearance that should be observed by the camera is uniform gray, and all projectors have the same, spatially uniform form factors. An object occludes only a single projector. A shadow appears from the left side in the captured image coordinate, and moves rightward at a constant velocity.

When no feedback structure is applied, the occluded pixels are

not compensated and remain dark; see Fig. 4(top). On the other hand, the distributed feedback algorithm is effective in this situation because it can manage the plug-out of a projector which is essentially the same situation as the emergence of shadow [TIK15]. The projection result recovers gradually because other occlusion-free projectors compensate for the shadow by the distributed feedback; see Fig. 4(middle). It should be noted that this phenomena is identical over all pixels because this feedback scheme is pixel-independent and reacts to the occlusion of each pixel. On the other hand, we achieve a more efficient shadow removal as shown in Fig. 4(bottom), where the shadow vanishes more quickly than our previous method [TIK15]. In the next section, we describe how the proposed method achieves such superior performance.

4. Proposed Algorithm

We propose an algorithm to remove undesirable shadows generated by moving occluders. From Sec. 4.1 to Sec. 4.3, we explain our proposed method assuming a simplified situation described in Sec. 3.3 where a moving object occludes a single projector. A shadow appears at the leftmost pixels including the pixel of $j = 1$ at time $k = 1$, and moves rightward at the velocity of 1 [pixel/frame] as shown in Fig. 4. Then, we formulate the proposed method in more general form in Sec. 4.4. We denote \mathcal{F} as the set of occlusion-free projectors ($\mathcal{F} = \{2, 3, \dots, n\}$ if only the projector 1 is subject to the occlusion).

4.1. Shadow removal via spatio-temporal prediction

We focus on a specific row of captured images, say the one including the pixel of $j = 1$. From the assumption mentioned above, a pixel j begins being occluded at time $k = j$ just before the observation. Let us define $s[j]$ by

$$s[j] = s[j-1] + \underbrace{\eta (r^j[j] - y^j[j])}_{e^j[j]}, \quad (12)$$

where η is a positive constant. $s[j]$ is physically the accumulation of error between the target and displayed intensities. Note that we consider the special case of $k = j$ to simplify the explanation of our proposed technique by assuming that the shadow moves at a constant velocity of 1 [pixel/frame]. General discussion is provided in Section 4.4.

We can calculate $s[j]$ when the projection error at the j -th pixel at time $k = j$ is observed. Then, we modify the broadcast data ξ in Eq. (5) as

$$\xi^{j+1}[j] = K e^{j+1}[j] + s[j], \quad (13)$$

only for a pixel where the shadow appears in the next frame, that is, $(j+1)$ -th pixel at time $k = j$ in this particular situation. The processing at the projector node is the same as Eq. (6). The distributed structure is retained in that the camera node does not require additional information. The first term of the right hand side is the pixel-wise feedback compensation as in the previous method. The newly added second term is aimed at cancelling the effect of the shadow. More specifically, when the shadow appears at the j -th pixel at time $k = j$, $y^j[j]$ becomes smaller, which leads to larger $e^j[j]$. We expect that similar phenomena (i.e., intensity reduction of the projected

result) happens at the $(j + 1)$ -th pixel at time $k = (j + 1)$ by the shadow's movement. To prevent this, Eq. (12) increases $s[j]$, which is added to $\xi^{j+1}[j]$. As a result $y^{j+1}[j + 1]$ increases, which reduces the performance deterioration caused by the shadow. Conceptually, the proposed method propagates the error information to the next frame to reduce the shadow effect at newly occluded pixels.

The recovery from the occlusion, where newly activated projectors generate unnecessary lights, can be handled in a similar manner. In that case, \mathcal{F} denotes the set of all the active projectors after the recovery ($\mathcal{F} = \{1, 2, \dots, n\}$ when the obstacle occludes no projector). The proposed method is hereinafter referred to as **error propagation** method.

4.1.0.1. Mathematical analysis: We conduct a mathematical analysis for linear dynamics ignoring the saturation \mathcal{P} and disturbance d . This also tells us how to determine the parameters K, η, ζ . For simplicity in the analysis, we further assume

- Pixel-wise feedback control achieves the perfect projection before the arrival of the occlusion, that is,

$$e^j[k] = 0 \text{ for } k < j. \quad (14)$$

- The spatial non-uniformity of the form factors of the occlusion-free projectors is ignored such that

$$\sum_{i \in \mathcal{F}} (p_i^j)^2 = \kappa_f, \forall j. \quad (15)$$

The occlusion of j -th pixel at the next frame (i.e., $k = j$) causes error

$$e_c^j := r^j - \sum_{i \in \mathcal{F}} p_i^j x_i^j[j - 1],$$

when no compensation is performed.

Let us consider the error at j -th pixel at time $k = j$. First, we have

$$e^j[j] = r^j - \sum_{i \in \mathcal{F}} p_i^j x_i^j[j] \quad (16)$$

$$= r^j - \sum_{i \in \mathcal{F}} p_i^j \left(x_i^j[j - 1] + p_i^j \xi^j[j - 1] \right) \quad (17)$$

$$= e_c^j - \kappa_f (K e^j[j - 1] + s[j - 1]) \quad (18)$$

$$= e_c^j - \kappa_f s[j - 1]. \quad (19)$$

By combining this with (12), we obtain

$$e^j[j] - e^{j-1}[j - 1] = (e_c^j - e_c^{j-1}) - \kappa_f (s[j - 1] - s[j - 2]) \quad (20)$$

$$= (e_c^j - e_c^{j-1}) - \kappa_f \eta e^{j-1}[j - 1] \quad (21)$$

and consequently,

$$e^j[j] = (1 - \kappa_f \eta) e^{j-1}[j - 1] + (e_c^j - e_c^{j-1}). \quad (22)$$

Therefore, the error system is stable if and only if

$$-1 < 1 - \kappa_f \eta < 1, \quad (23)$$

or equivalently

$$0 < \kappa_f \eta < 2.$$

Actually, if the projection profile is spatially uniform such that

$$e_c^j = \bar{e}_c \text{ for all } j, \quad (24)$$

$e^j[j]$ converges to 0 as j increases, independent of \bar{e}_c . This is why the shadow effect is unobservable for $k = 20$ in Fig. 4(bottom). More generally, considering the effect of the previous pixel-wise feedback, we have

$$e^j[k] = (1 - \kappa_f \eta)^{j-1} (1 - \kappa_f K)^{k-j} \bar{e}_c \quad (25)$$

for $k \leq j$. This explains the whole dynamical behavior in Fig. 4; (top) for $K = \eta = 0$, (middle) for $K > 0$ and $\eta = 0$, and (bottom) for $K > 0$ and $\eta > 0$.

The assumption Eq. (24) is not essential. Actually, the performance for general case can also be analyzed based on simple spatio-temporal frequency response calculus. This is omitted because it seems beyond the main interest of an audience.

4.2. Role of the servo mechanism

Some remarks on the structure of the proposed algorithm are given in this section. Again, the projection profile before the occlusion is assumed to be spatially uniform such that

$$e_c^j = \bar{e}_c, \forall j.$$

It should be noted that perfect shadow removal is achieved if

$$s[j] = s^* := \frac{\bar{e}_c}{\kappa_f}$$

because this completely cancels the effect of the occlusion through the broadcast and projection update; see Eq. (19). However, $s[j]$ should be calculated at the camera node where the form factor data κ_f is unavailable. This is the essential difficulty of the shadow removal in the distributed environment in this paper.

To prevent this issue, we utilized the servo mechanism that can track arbitrary step signals. This is implemented in the update of $s[j]$ in an accumulation manner. As a result, as far as η satisfies Eq. (23), we obtain

$$s[j] \rightarrow s^* \text{ as } j \rightarrow \infty$$

independent of η and \bar{e}_c . This is the reason why we successfully achieved *asymptotically perfect shadow removal* without the information of κ_f .

One may think that a naïve algorithm

$$s[j] = \bar{\eta} e^j[j] \quad (26)$$

can provide a similar effect. In this case, the result is very sensitive to the choice of $\bar{\eta}$, and stationary shadow remains although it is reduced; see Fig. 5(a) where $\bar{e}_c = 0.3$ and $\kappa_f = 0.625$. It can readily be verified that $\kappa_f \bar{\eta} < 1$ is required for the stability, and that the stationary error is equal to $\bar{e}_c / (1 + \kappa_f \bar{\eta})$. This clarifies the performance limitation of Eq. (26) in that 50% removal is not possible. This shows a clear contrast to the asymptotic perfect shadow removal of the proposed algorithm whose performance is shown in Fig 4(bottom). Fig. 5(b) shows the dynamical behavior of $e^j[j]$.

4.3. Effect of saturation and modified algorithm

In Eq. (12), the effect of shadow that should be compensated for in the next pixel is measured by the error $e^j[j]$. However, when there is

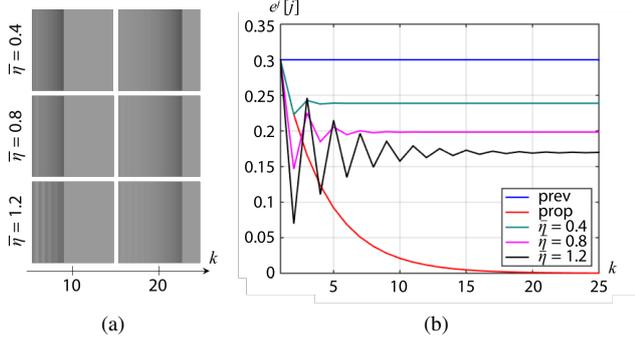


Figure 5: Simulation of naïve algorithm in Sec. 4.2: (a) simulated shadow removal results when Eq. 27 is applied with various $\bar{\eta}$ under the same simulation condition of Fig.4, (b) dynamical behavior of $e^j[j]$ (prev: previous method, prop: proposed method).

some stationary error due to the saturation, $e^j[j]$ cannot necessarily capture the effect of shadow. Instead, we can implement

$$s[j] = s[j-1] + \eta(y^j[j] - y^j[j-1]) \quad (27)$$

where the difference between the projection result before/after occlusion is regarded as the effect of the occlusion. This seems to be more suitable for the measure of the shadow effect because this is free from the stationary error at the cost of the memory for $y^j[j-1]$. This modified method is hereinafter referred to as **change propagation method**.

For linear analysis similar to Sec. 4.1, we can combine Eq. (20) with Eq. (27) to have

$$e^j[j] - e^{j-1}[j-1] = (e_c^j - e_c^{j-1}) - \kappa_f(s[j-1] - s[j-2]) \quad (28)$$

$$= (e_c^j - e_c^{j-1}) - \kappa_f \eta (e^{j-1}[j-1] - e^{j-1}[j-2]) \quad (29)$$

$$= (e_c^j - e_c^{j-1}) - \kappa_f \eta e^{j-1}[j-1], \quad (30)$$

where we utilized $y^{j-1}[j-1] - y^{j-1}[j-2] = e^{j-1}[j-1] - e^{j-1}[j-2]$. Therefore, we obtain identical equation to Eq. (22) in this simplified analysis. This means that error propagation and change propagation produce the same projection result as far as Eq. (14) is satisfied.

4.4. General algorithm

Instead of the simplified situation discussed above, we formulate our proposed algorithms for more general situation, which will be implemented in our experiments.

4.4.0.1. Error propagation method: Let $\mathcal{O}[k]$ denotes the set of newly occluded pixels at time k ($\mathcal{O}[k] = \{k\}$ in the previous simplified case). In this general case, we can replace Eq. (12) by

$$s[k] = s[k-1] + \eta \frac{1}{|\mathcal{O}[k]|} \sum_{j \in \mathcal{O}[k]} e^j[k], \quad (31)$$

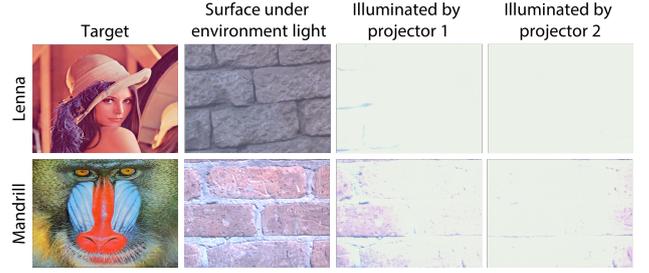


Figure 6: Target images and projection surfaces: (let column) target image, (second column) the appearance of projection surface under environment light, (third column) the surface under uniform white illumination (i.e., form factor) by projector 1, and (fourth column) that by projector 2.

where $|\mathcal{O}[k]|$ is the number of pixels contained in $\mathcal{O}[k]$. Then, we replace Eq. (13) by

$$\xi^j[k] = Ke^j[k] + s[k] \text{ for } j \in \mathcal{O}[k+1]. \quad (32)$$

For the pixels which do not belong to $\mathcal{O}[k+1]$, the broadcast data $\xi^j[k]$ is computed using Eq. (5).

4.4.0.2. Change propagation method: In the change propagation method, we can replace Eq. (27) by

$$s[k] = s[k-1] + \eta \frac{1}{|\mathcal{O}[k]|} \sum_{j \in \mathcal{O}[k]} (y^j[k] - y^j[k-1]), \quad (33)$$

which is then used to compute the broadcast data using Eq. (32).

5. Experiment

We conducted both physical and simulation experiments to validate the proposed algorithm.

5.1. Physical experiment

We built a projector-camera system consisting of two projectors (Epson EH-TW410) and one camera (PointGrey Flea3 FL3-U3-88S2C-C, 1600×1200 pixels). Figure 1(left) shows an overview of the whole system. During the projection, a rectangular board, which is attached on a locomotion robot (LEGO® MindStorm EV3) as in Fig. 1(left), started to move at time $k=0$ from left to right in front of a projection surface at a almost constant speed with the average velocity of 16 [mm/s]. The camera node and projector nodes were implemented on a single laptop computer (Sony VAIO-Duo 11 SVD1123AJ, CPU: Core i7 3687U 2.10GHz 2.60GHz, RAM: 8.0GB), which were virtually separated into independent systems. Each (virtual) projector node does not have any information about the occlusion.

We assume that the robot movement was the same in every trial through the experiment. Each trial took 50 frames. We did not predict shadow regions during runtime. Instead, we approximated the shape of a shadow region as a rectangle, and manually measured its velocity in the camera coordinate system in advance. The measured average velocity was 250 [pixel/frame].

We prepared two experimental conditions, **lenna** and **mandrill**, as the combinations of different target images and projection surfaces as shown in Fig. 6. Two images, lenna and mandrill, were applied as the target images. The projection surfaces were planar boards, on which different brick textures were printed. The form factors are also shown in Fig. 6.

From Eq. (25), the optimal parameter values for η and K are naturally $1/\kappa_f$. However, it is impossible to obtain exact value of κ_f due to measurement errors. In addition, it is desirable from an implementation viewpoint that a common κ_f is used for all pixels. A histogram of the form factors of both the projectors shown in Fig. 6 is given by Figure 7. Considering this distribution and Eq. (15), let us regard $\kappa_f = 1.7$ by taking the nominal value of R channel which is middle among the 3 color channels.

5.1.1. Shadow removal performance

We compared the shadow removal performances among the previous method [TIK15], the proposed error propagation method, and the change propagation method. Because the robot moved at almost constant speed, we applied a fixed value 250 [pixel/frame] (i.e., the measured average velocity) as the velocity of the shadow region movement v to predict the positions of shadow pixels in the next frame. K and η took 0.6 as $1/\kappa_f$.

Figure 1(right) shows the projected result sequences in the lenna condition. As shown in the figure, we confirmed that the proposed methods (both error propagation and change propagation) provided better shadow removal performances. Figure 8(a) shows the time series of the mean square errors (MSEs) in both the lenna and mandrill conditions. Note that the MSEs were computed using 8-bit image data. A one-way analysis of variance (ANOVA) with repeated measures showed statistically significant differences among the MSEs of all the methods in the lenna condition ($p < 0.01$) and the mandrill condition ($p < 0.01$). Post-hoc analysis was then performed using Ryan’s method for pairwise comparison. It showed statistically significant differences between the previous and error propagation methods, and the previous and change propagation

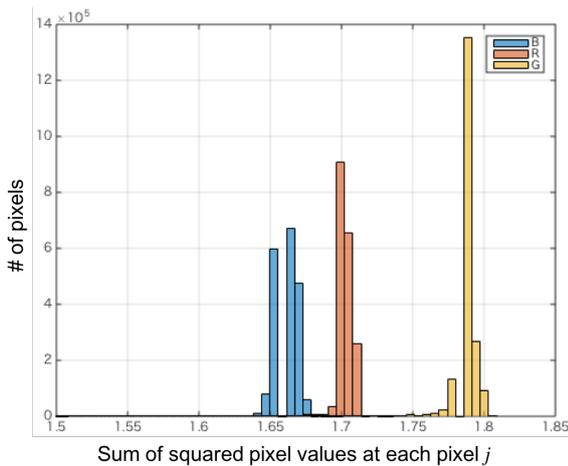


Figure 7: Histogram of $\sum_{i \in \mathcal{F}} (p_i^j)^2$, $\mathcal{F} = \{1, 2\}$.

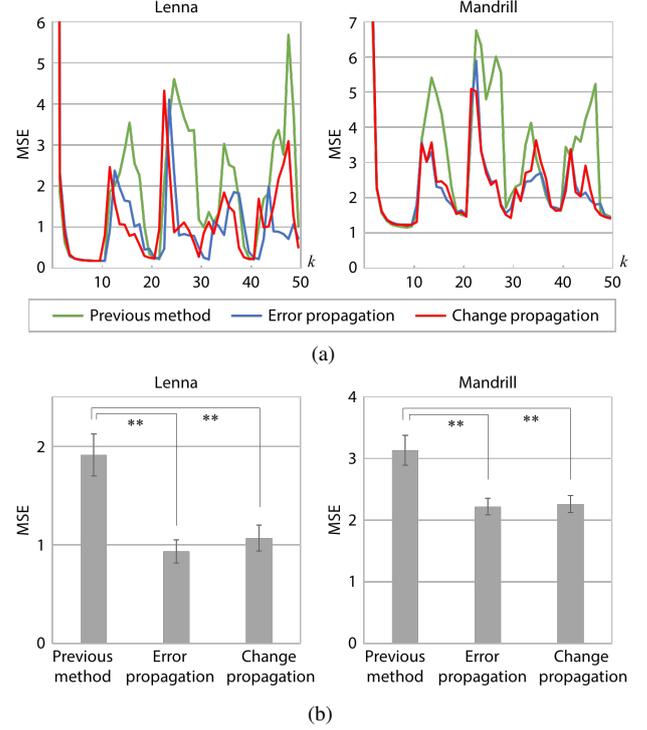


Figure 8: Mean squared error (MSE): (a) the time series of MSEs, (b) the average with standard error of MSEs of $6 \leq k \leq 50$ (**: $p < 0.01$).



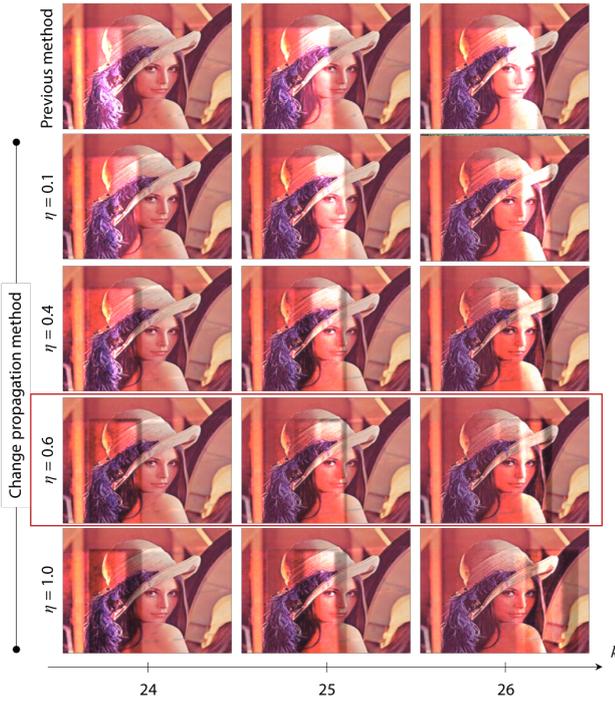
Figure 9: Broadcasted data ξ .

methods in both the lenna and mandrill conditions ($p < 0.01$) as shown in Fig. 8(b). From this result, we quantitatively confirmed that both the proposed methods could display the target images with significantly less errors than the previous method in both conditions. Therefore, we regard the parameter values of $K = 0.6$, $\eta = 0.6$, $v = 250$ were suitable for this experimental setup.

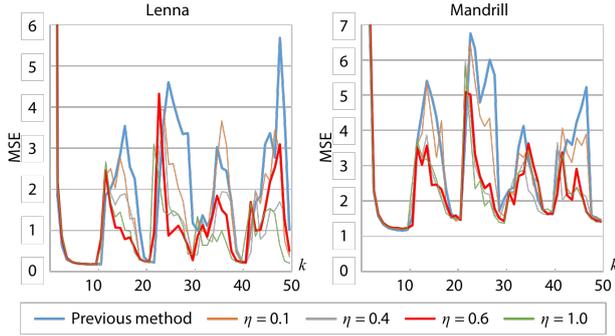
Figure 9 shows the broadcasted data ξ at the 15th frame. We could see the propagated information indicated by the yellow arrows. Thanks to this error/change propagation, the proposed methods outperformed the previous method.

5.1.2. Parameter robustness

In this section, we evaluated the parameter robustness by conducting the same shadow removal experiments using the change propagation method with different η and v values from the suitable ones (i.e., $\eta = 0.6$, $v = 250$).



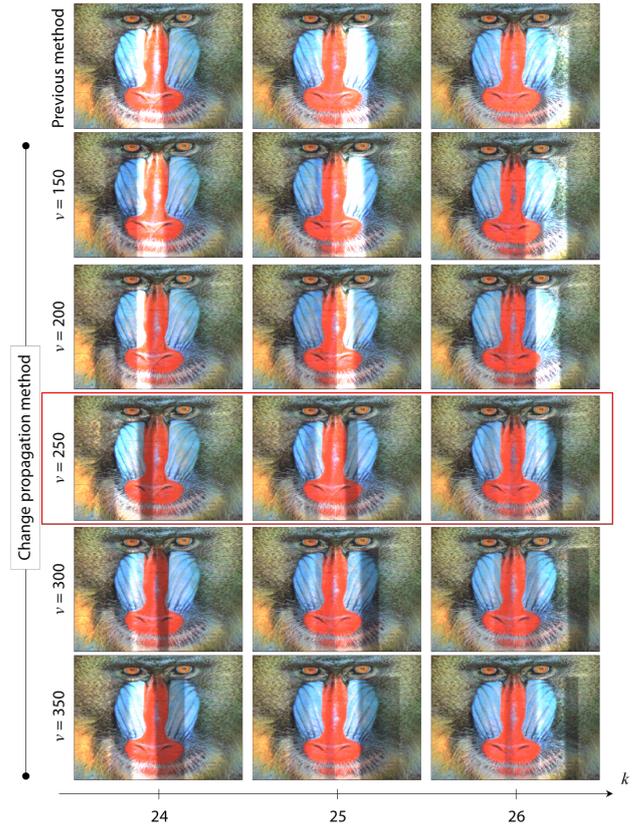
(a)



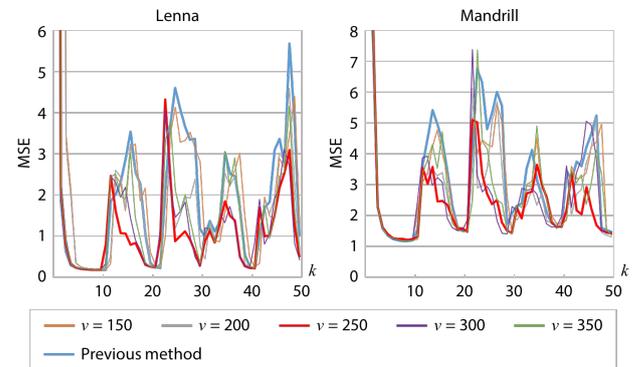
(b)

Figure 10: Experimental results of the change propagation method with different η and the previous method: (a) projected results at $k = 24, 25, 26$ in the lenna condition (the red frame indicates the results with the suitable value of η), (b) the time series of MSEs.

5.1.2.1. Evaluation of η : We conducted the shadow removal experiments with different η values while the other parameter values were not changed from the suitable ones. In particular, we applied the following values: $\eta = 0.1, 0.4, 0.6, 1.0$, where 0.6 was the suitable one. Figure 10(a) shows the projected result sequences of the previous method and the change propagation methods with different η values. Figure 10(b) shows the time series of MSEs. From the results, we confirmed that the proposed system with different η values provided better shadow removal results than the previous method, while the improvement is not so significant with the smaller η value (i.e., $\eta = 0.1$).



(a)



(b)

Figure 11: Experimental results of the change propagation method with different v : (a) projected results at $k = 24, 25, 26$ in the mandrill condition (the red frame indicates the results with the measured value of v), (b) the time series of MSEs.

5.1.2.2. Evaluation of v : We conducted the shadow removal experiments with different v values while the other parameter values were not changed from the suitable ones. In particular, we applied the following values: $v = 150, 200, 250, 300, 350$ [pixel/frame], where 250 was the measured average velocity. Figure 11(a) shows the projected result sequences of the change propagation methods

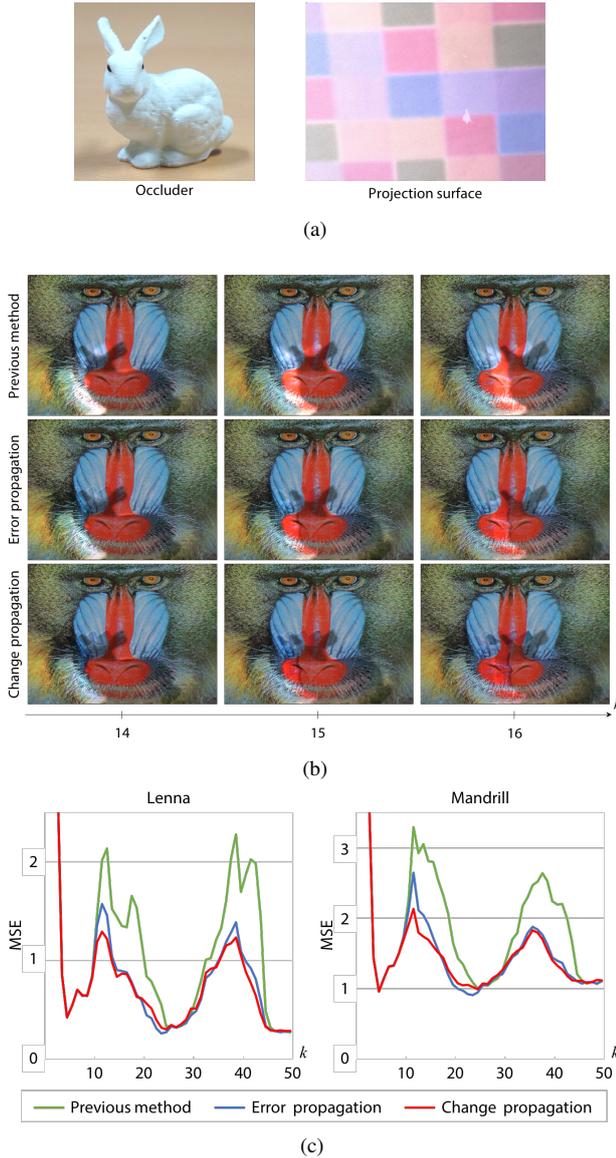


Figure 12: Physical experiment with a complex occluder: (a) the occluder and projection surface, (b) projected results (contrast enhanced for easier comparison), (c) the time series of MSEs.

with different v values. Figure 11(b) shows the time series of MSEs. We found that the errors were reduced from the previous method even when applied velocity values were different from the measured one. Taking a more in-depth look at the results in Fig. 11(b), we could find the velocity value of $v = 300$ provided results very close to those with the measured velocity value $v = 250$. From this evaluation, we confirmed that the shadow velocity v should be predicted during runtime, which however should not be perfectly accurate. More particularly, the prediction error was more acceptable when it was smaller and when velocity was overestimated than underestimated. We discuss the shadow velocity (or the next shadow area) estimation in Sec. 6.

5.1.3. Complex Occluder

We conducted another experiment with a complex shape occluder that is a Stanford bunny fabricated from a 3D printer (Fig. 12(a)). For this additional experiment, we used different equipment. In particular, a PC (CPU: Intel Core i7-960 3.20 GHz, RAM: 12 GB) controlled two projectors (Epson EMP-1710 and EMP-1715), a camera (PointGrey Flea3 FL3-U3-13S2C, 1280×960 pixels), and a slide stage (SUS XA-50L-600E) that moved the occluder. We prepared a projection surface whose texture was shown in Fig. 12(a). We applied the same target images (lenna and mandrill) as the previous experiment. Note that their contrasts were adjusted in advance to fit the limited dynamic range of the projectors. The procedure of the experiment was same as the previous one.

Figure 12(b) shows the projected result sequences of the previous, error propagation, and change propagation methods. Figure 12(c) shows the time series of MSEs, in which errors in the proposed methods (error propagation and change propagation) are clearly smaller than those in the previous method. From this evaluation, we confirmed that the proposed methods provided better compensation qualities than the previous method even when an occluder’s shape was relatively complex.

5.2. Simulation experiment for complex shadow movement

We conducted a simulation experiment to validate our proposed methods for more complex shadow movement. A natural image, whose image size was 512×512 pixels, was used as a target appearance as shown in Fig. 13(a). We captured the sequence of silhouette of a waving human hand as shadow regions of the experiment as shown in Fig. 13(b, top row). Therefore, we could test the shadow removal methods with temporally varying moving directions and speeds. In this experiment, shadow regions were known in advance, which were used in computing Eqs. 31 and 33.

Figures 13(b) and (c) show the simulation results. From Fig. 13(b), we could find that both the proposed methods provided better shadow removal appearances for both horizontal and vertical movements than the previous method. The graph shown in Fig. 13(c) quantitatively showed that the proposed methods could remove shadows with less errors than the previous method. Therefore, we confirmed that the proposed method worked for a moving shadow that changes its moving direction and speed during runtime.

6. Discussion

The proposed methods (error propagation and change propagation) provided faster shadow removal with less errors than the previous feedback method. We would like to emphasize again that the proposed methods are designed in a distributed manner, where there is no centralized controller that computes projection images for all the projectors. Broadcast from the camera node to the projector nodes is the only one communication channel in the system. In this paper, we proposed to modify the previous distributed optimization method by adding error or change propagation information to the

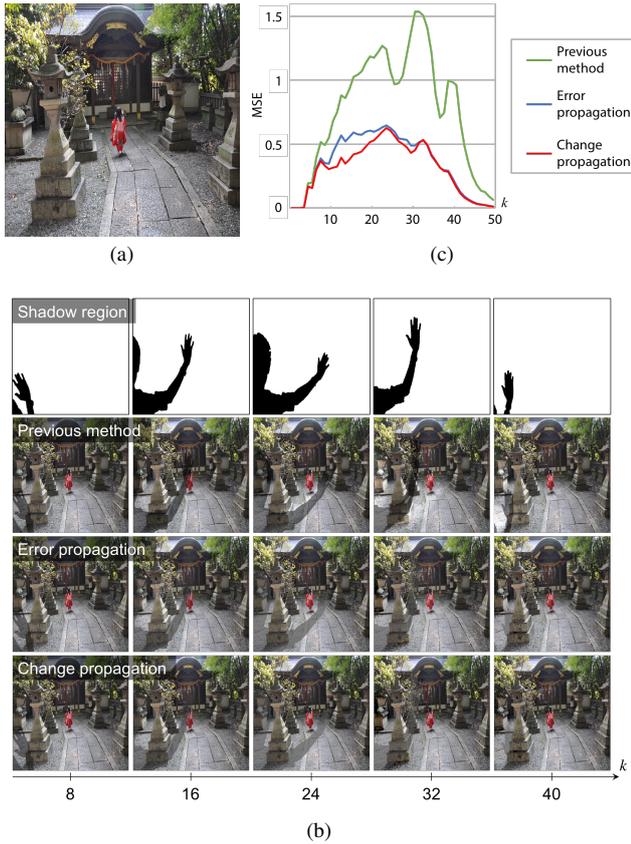


Figure 13: Simulation experiment: (a) target image, (b) shadow region and simulated result, (c) the time series of MSEs.

broadcasting data. In particular, the error/change propagation information, computed from pixels in the current shadow region, is added to the predicted next shadow region of the broadcasting data.

Although shadow prediction is not the focus of our paper, it is needed in our method as discussed above. In Sec. 5.1.2, we confirmed that the prediction of the next shadow regions does not have to be perfect, but should be accurate to some extent. For example, in our experimental condition where shadow moved in one direction at a constant speed, the proposed method worked well with prediction errors of 20 % of the ground truth (error: 50 pixels, ground truth: 250 pixels). Shadow prediction with this level of errors can be integrated with our method by applying the combination of a shadow region detection such as [SFC*07] and an optical flow detection [Far03]. We believe that it is also possible to detect shadow regions for free form objects (e.g., human body) by applying a deformable object tracking technique using the contour information such as [RVTY05]. Once it is integrated, our method can remove moving shadows whose moving directions and velocities change during runtime, as shown in our simulation experiment (Sec. 5.2)

The current system runs at around 2 fps on a laptop computer without any optimizations. Because most parts of the proposed methods can be easily parallelized and implemented on GPU, it

is technically possible to run the proposed methods in real-time. Decreasing the processing time generally leads to a better shadow prediction, and consequently, improves the shadow removal performance of the proposed methods.

It should be noted that the proposed methods also work for plug-in and plug-out situations. Although we did not explicitly demonstrate them in our experiments, shadow removal process shown in our experiments essentially contains plug-in and plug-out. In particular, shadow removal is equivalent to plug-out, because in both cases contributions from some projectors in some areas are lost. On the other hand, plug-in, in which newly added projectors start to project images on a surface, is equivalent to a situation where occluded projectors that become visible from a shadow area resume projecting images onto it.

7. Conclusion

We proposed a novel shadow removal algorithm based on the spatio-temporal prediction of the shadow regions. The effectiveness and robustness were examined through mathematical analysis and experiments. The most important contribution of this work is that we achieved an efficient shadow removal even in a distributed projector-camera system, in which only broadcasting from the camera node to the projector nodes is available. Thus, we realized a highly scalable and robust shadow removal compared to previous methods that assumed centralized control over all the projector nodes.

As noted in Sec. 1, we developed the proposed technique assuming a situation where multiple projection surfaces are located at different depths. For capturing occlusion-free images in such situation, we plan to apply synthetic aperture capturing technique [INS14], which can be implemented as a simple extension of the system proposed in this paper, and would be one of the future works. As another future work, we will integrate a shadow prediction with the proposed method, and implement the algorithm on GPU to make a system running in real-time.

Acknowledgements

This research was supported by JSPS KAKENHI Grant Number 26630198 and 15H05925.

References

- [AC07] AUDET S., COOPERSTOCK J. R.: Shadow removal in front projection environments using object tracking. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (2007), pp. 1–8. 2
- [AK10] AMANO T., KATO H.: Appearance control using projection with model predictive control. In *Proceedings of International Conference on Pattern Recognition (ICPR)* (2010), pp. 2832–2835. 2
- [ASUK14] AMANO T., SHIMANA I., USHIDA S., KONO K.: Successive wide viewing angle appearance manipulation with dual projector camera systems. In *Proceedings of International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments (ICAT-EGVE)* (2014), pp. 49–54. 2
- [BEK05] BIMBER O., EMMERLING A., KLEMMER T.: Embedded entertainment with smart projectors. *IEEE Computer* 38, 1 (2005), 56–63. 3

- [BIWG08] BIMBER O., IWAI D., WETZSTEIN G., GRUNDHÖFER A.: The Visual Computing of Projector-Camera Systems. *Computer Graphics Forum* 27, 8 (2008), 2219–2254. 2
- [BR05] BIMBER O., RASKAR R.: *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A. K. Peters Ltd., 2005. 2
- [Far03] FARNEBÄCK G.: Two-frame motion estimation based on polynomial expansion. In *Proceedings of the 13th Scandinavian Conference on Image Analysis* (2003), pp. 363–370. 4, 10
- [GPNB04] GROSSBERG M. D., PERI H., NAYAR S. K., BELHUMEUR P. N.: Making one object look like another: controlling appearance using a projector-camera system. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on* (June 2004), vol. 1, pp. I-452–I-459 Vol.1. doi:10.1109/CVPR.2004.1315067. 3
- [INS14] IWAI D., NAGASE M., SATO K.: Shadow removal of projected imagery by occluder shape measurement in a multiple overlapping projection system. *Virtual Reality* 18, 4 (2014), 245–254. 2, 10
- [JWS*01] JAYNES C., WEBB S., STEELE R. M., BROWN M., SEALES W. B.: Dynamic shadow removal from front projection displays. In *Proc. IEEE Visualization* (2001), pp. 175–182. 2
- [JWS04] JAYNES C., WEBB S., STEELE R. M.: Camera-based detection and removal of shadows from interactive multiprojector displays. *IEEE Transactions on Visualization and Computer Graphics* 10, 3 (2004), 290–301. 2
- [NO09] NEDIC A., OZDAGLAR A.: Distributed Subgradient Method for Multi-agent Optimization. *IEEE Transactions on Automatic Control* 54, 1 (2009), 48–61. 2
- [RVTY05] RATHI Y., VASWANI N., TANNENBAUM A., YEZZI A.: Particle filtering for geometric active contours with application to tracking moving and deforming objects. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (June 2005), vol. 2, pp. 2–9 vol. 2. doi:10.1109/CVPR.2005.271. 10
- [SCS01] SUKTHANKAR R., CHAM T.-J., SUKTHANKAR G.: Dynamic shadow elimination for multi-projector displays. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (2001), vol. II, pp. 151–157. 2
- [SFC*07] SUMMET J., FLAGG M., CHAM T.-J., REHG J. M., SUKTHANKAR R.: Shadow elimination and blinding light suppression for interactive projected displays. *IEEE Transactions on Visualization and Computer Graphics* 13, 3 (2007), 508–517. 2, 10
- [SMK10] SUGAYA Y., MIYAGAWA I., KOIKE H.: Contrasting shadow for occluder light suppression from one-shot image. In *Proc. International Workshop on Projector-Camera Systems* (2010), pp. 13–18. 2
- [Ste15] STEPHEN BOYD.: http://www.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf, 2015. 3
- [TIK15] TSUKAMOTO J., IWAI D., KASHIMA K.: Radiometric compensation for cooperative distributed multi-projection system through 2-dof distributed control. *IEEE Transactions on Visualization and Computer Graphics* 21, 11 (2015), 1221–1229. 1, 2, 3, 4, 7