Visual Feedback Without Geometric Features Against Occlusion: A Walsh Basis

Satoru Sakai, *Member, IEEE*, Masayuki Ando, and Shunsuke Kobashi

***Abstract*— For a visual feedback without geometric features, this brief suggests to apply a basis made by the Walsh functions in order to reduce the off-line experimental cost. Depending on the resolution, the feedback is implementable and achieves the closed-loop stability of dynamical systems as long as the input–output linearity on matrix space exists. Remarkably, a part of the whole occlusion effects is rejected, and the remaining part is attenuated. The validity is confirmed by the experimental feedback for nonplanar sloshing.**

***Index Terms*— Dynamical systems, occlusion, stability, visual feedback.**

I. INTRODUCTION

IN MANY conventional visual feedbacks, there exists a series interconnection between the control block and the image processing block in Fig. 1. In the image processing block, the geometric features (e.g., a dot position and a line angle) are defined and extracted from the camera image on line. Via the series interconnection, a lot of information is lost in the image processing block, but the design procedures of the control block can be systematic when fruitful control theories are applicable. On the other hand, the design procedures of the image processing block are not or less systematic, especially in the presence of occlusion (visual obstacles between the camera and the object), because the way to define and extract geometric features strongly depends on the plant block, the control objective, and so on.

To solve this problem, not many but several visual feedbacks without or with less geometric features are discussed by different approaches, such as the homography-based approach [1], [2] and the Hausdorff distance-based approach [3]. The similar motivation is traced back to the subspace approaches [4], [5]. Most of them could work locally at least for static systems that are acceptable when the camera or object dynamics (e.g., the camera-link flexibility) are negligible. On the other hand, the closed-loop stability of dynamical systems is not guaranteed and can be lost even in the absence of occlusion. Exceptionally, a visual feedback [6], [7] locally guarantees the closed-loop stability of a special nonlinear dynamical system assuming the absence of occlusion.

In this brief, in the presence of occlusion, a visual feedback without geometric features is given as a new application for linear dynamical systems. The closed-loop stability is guaranteed by our simple idea beginning with a change of the mapping domain and codomain (the input and output spaces) of the plant block. In many conventional visual feedbacks, geometric features are defined in a coordinate space $\mathbb{R}^n$ (e.g., the camera image plane $\mathbb{R}^2$), which can be eventually the output space of the plant block. In our visual feedback in Fig. 2, geometric features are not defined, and a matrix space $\mathbb{M}^{m\times n}$ is the output space of the plant block. Since any coordinate space is isomorphic to a matrix space, the design procedures of our control block on matrix space can be systematic when fruitful control theories are applicable again.

However, due to the computational limitation at least, such theories are not always applicable as they are. In our visual feedback, from the perspective of the Hilbert space [8], we can design a basis in the output space $\mathbb{M}^{m\times n}$ so that the control theories are applicable under the computational limitation. Indeed, in the absence of occlusion, our pilot study [9] performs an off-line basis generation procedure before the system identification procedure.

In the presence of occlusion, this brief suggests to apply a new special basis by which any off-line basis generation procedure is not needed. This means a cut of the experimental cost, because the experimental movies for the off-line basis generation procedure are nothing but big data for control. The new special basis is made by the Walsh functions, which have not been applied for modeling and control of dynamical systems by the conventional visual feedbacks without geometric features.

The rest of this brief is organized as follows. In Section II, dynamical systems on matrix space are introduced, and the new special basis is suggested for our visual feedback.

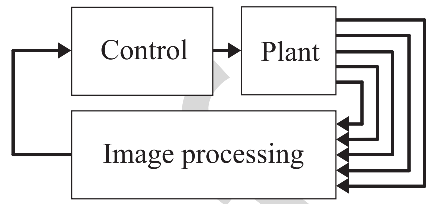

Fig. 1. Block diagram with the image processing for geometric features.

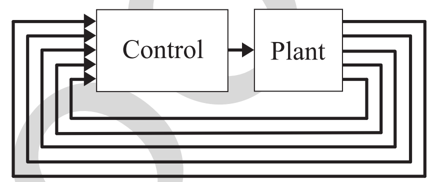

Fig. 2. Block diagram without the image processing for geometric features.

Manuscript received January 24, 2017; revised October 2, 2017; accepted November 30, 2017. Manuscript received in final form December 1, 2017. This work was supported by JSPS KAKENHI under Grant 17K06227. Recommended by Associate Editor A. Behal. *(Corresponding author: Satoru Sakai.)*

The authors are with the Department of Mechanical Engineering, Shinshu University, Matsumoto 4-17-1, Japan (e-mail: satorusakai@gakushikai.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCST.2017.2780176

The new special basis does not need any off-line basis generation procedure but can be systematically truncated without geometric features so that fruitful control theories are applicable under the computational limitation. In Section III, the proposed control is applied to nonplanar sloshing whose dynamics is not negligible. The validity is confirmed experimentally in the presence of occlusion. Finally, this brief is concluded in Section IV.

II. Dynamical Systems on Matrix Space

Let us consider a finite-dimensional space denoted by $\mathbb{M}^{m\times n}$ of a time-varying matrix $M(k) \in \mathbb{M}^{m\times n}$ at the discrete-time instant $k \in \mathbb{Z}_+ := \{0\}\cup\mathbb{N}$. The matrix space $\mathbb{M}^{m\times n}$ is a Hilbert space with the inner product

$$\langle M(k), N(k)\rangle = \mathrm{tr}(M(k)^{\mathrm{T}} N(k)) \in \mathbb{R}$$

for all matrices $M(k)$ and $N(k) \in \mathbb{M}^{m\times n}$. $M(k) \perp N(k)$ implies $\langle M(k), N(k)\rangle = 0$, and the inner product introduces the norm $\|M(k)\| = (\langle M(k), M(k)\rangle)^{1/2} \geq 0$. The notation $\mathrm{tr}(\bullet)$ denotes the trace of a matrix. Consider a finite-dimensional linear time-invariant (LTI) system described by linear mappings between matrix spaces [9]

$$\begin{cases} X(k+1) = \mathcal{A}\circ X(k) + \mathcal{B}\circ U(k) + V(k) \\ Y(k) = \mathcal{C}\circ X(k) + \mathcal{D}\circ U(k) + W(k) \end{cases} \tag{1}$$

where the state $X(k) \in \mathbb{M}^{m_x\times n_x}$ and the state disturbance $V(k) \in \mathbb{M}^{m_x\times n_x}$ are the $m_x \times n_x$ matrices, the input $U(k) \in \mathbb{M}^{1\times 1}$ is the 1×1 matrix, and the output $Y(k) \in \mathbb{M}^{m_y\times n_y}$ and the output disturbance $W(k) \in \mathbb{M}^{m_y\times n_y}$ are the $m_y \times n_y$ matrices. The notation $\circ$ denotes the operation of the linear mappings $\mathcal{A}, \mathcal{B}, \mathcal{C}$, and $\mathcal{D}$.

Remark 1: Since every mapping cannot be defined until the domain and the codomain are defined, every system depends on the choice of the input and output spaces. In this sense, the proposed system (1) choosing the matrix spaces and the well-known LTI system choosing the coordinates spaces are different mathematical objects even if the linear mappings $\mathcal{A}, \mathcal{B}, \mathcal{C}$, and $\mathcal{D}$ of both systems have the same matrix representations. On the other hand, since there is an isomorphism from a matrix space $\mathbb{M}^{m\times n}$ to a coordinate space $\mathbb{R}^{mn}$ [8], fruitful control theories [e.g., ARX, N4SID, linear quadratic and Gaussian (LQG), and μ] are applicable to both systems.

Remark 2: The output Y corresponding to the camera image is visible, but the input U and the state X are invisible as they are the disturbances V and W. Of course, the input U is not unknown and visualizable, but the state X is not always visualizable even in the absence of the disturbances V and W.

It is never our contribution to see the camera image as a matrix and is popular in the image processing blocks that are regarded as static open systems. In our visual feedback, from the viewpoint of dynamical closed-loop systems including the plant block, not only the camera image corresponding to the output of the plant block, but also the input and state are matrices. The inner product (or the passivity) of the output Y and the input U can be taken when they belong to the same subspace. In mathematics, roughly speaking, a matrix space is almost the same as a coordinate space, which is familiar. However, in engineering, as long as the control objective is defined in the camera image, the matrix space is more suitable to design the basis.

Since a matrix space $\mathbb{M}^{m\times n}$ has a normalized orthogonal basis $E_1, \ldots, E_{mn}$ [8]

$$\langle E_{\ell_i}, E_{\ell_j}\rangle = \begin{cases} 0 \ (\ell_i \neq \ell_j) \\ 1 \ (\ell_i = \ell_j), \end{cases} \qquad \ell_i, \ell_j = 1, \ldots, mn$$

every time-varying matrix

$$M(k) = \sum_{\ell=1}^{mn} \langle M(k), E_\ell\rangle E_\ell \in \mathbb{M}^{m\times n}, \quad \ell = 1, \ldots, mn$$

has a representation $[m_1(k), m_2(k), \ldots, m_{mn}(k)]^{\mathrm{T}}$ whose component is of the form

$$m_\ell(k) := \langle M(k), E_\ell\rangle. \tag{2}$$

Here, the most popular basis in the output space is the standard basis (the pixel-by-pixel basis)

$$E_1^S = \begin{bmatrix} 1 \cdots 0 \\ \vdots \quad \vdots \\ 0 \cdots 0 \end{bmatrix}$$

$$E_2^S = \begin{bmatrix} 0 \ 1 \cdots 0 \\ \vdots \ \vdots \quad \vdots \\ 0 \ 0 \cdots 0 \end{bmatrix}, \ldots, E_{mn}^S = \begin{bmatrix} 0 \cdots 0 \\ \vdots \quad \vdots \\ 0 \cdots 1 \end{bmatrix}$$

by which any off-line basis generation procedure is not needed. The standard basis could work locally at least for static systems as the pixel-by-pixel feedback. However, the standard basis can cause several problems for dynamical systems. One of them is from the computational limitation, because the number of the standard basis elements is nothing but the number of the pixels mn, which is usually quite large [10]. Indeed, a more than 1×10^6 pixels feedback is implemented on a better hardware [2]. Nevertheless, the standard basis cannot be truncated systematically without geometric features. For example, for a certain plant block with a control objective, even if we know that the (1,2)-pixel of the camera image is not important, the truncation of E_2^S is not accepted, because such truncation is nothing but the geometric feature extraction depending on the plant block or the control objectives.

To solve the standard basis problem, under the computational limitation, our pilot study [9] discusses an alternative basis, which is systematically truncated without geometric features. However, the alternative basis needs an off-line basis generation procedure before the system identification procedure. This means an increase of the experimental cost, since the alternative basis cannot be generated without acquiring the experimental movies.

One may think that the experimental cost in the off-line basis generation procedure is not an issue, since the acquired movies for the off-line basis generation procedure can be reused for the system identification procedure. This is not true. The acquired movies for the off-line basis generation procedure are nothing but big data for control (e.g., the raw movies) and are much bigger than the outputs for the system

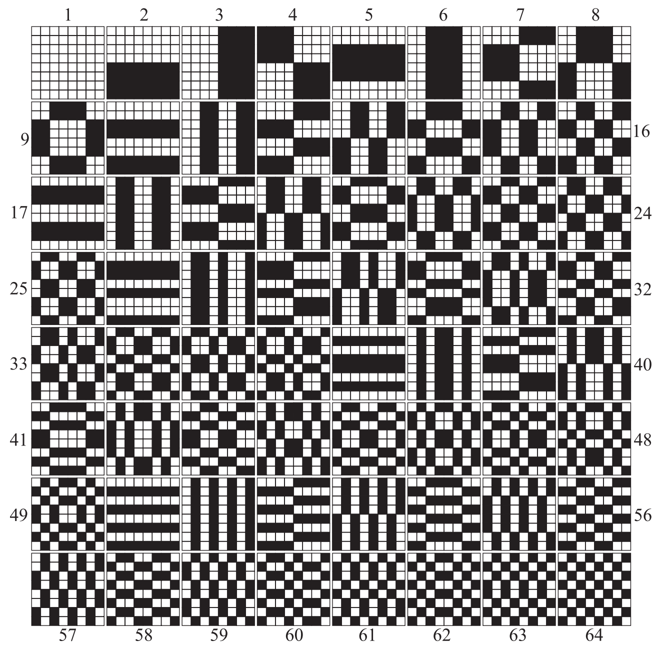

Fig. 3. Walsh basis in the order of the horizontal first and the vertical second sequence (white $= +1/64$ and black $= -1/64$).

identification procedure in which the number of the basis elements (the output dimension) is already fixed.

To solve the alternative basis problem as well as the standard basis problem, this brief suggests to apply a new special basis, which can be systematically truncated without geometric features but does not need any off-line basis generation procedure. Fig. 3 shows an example of the new special basis

$$E_\ell^W = \mathrm{vec}^{-1}\left(\left[W\left(\ell-1, \frac{0}{mn}\right)\cdots W\left(\ell-1, \frac{mn-1}{mn}\right)\right]^{\mathrm{T}}\right)$$

with

$$\underbrace{W(\ell-1, x) = (-1)^{\Sigma_{j=1}^{\infty} x_j (\ell-1)_{1-j}}}_{\text{the Walsh function [11]}}, \quad \ell = 1, \ldots, mn$$

whose $x_j \in \{0, 1\}$ and $(\ell-1)_{1-j} \in \{0, 1\}$ are determined by the dyadic expansion of the normalized space parameter $x = \Sigma_{j=1}^{\infty} x_j \cdot 2^{-j} \in [0, 1)$ and that of the number $(\ell - 1) = \Sigma_{j=0}^{K} (\ell-1)_{-j} \cdot 2^j \in \mathbb{Z}_+$ $(K \in \mathbb{Z}_+, (\ell-1) \in [2^K, 2^{K+1}))$. Here, the number $N := mn$ is constrained to be $m = n = 2^L$ $(\forall L \in \mathbb{Z}_+)$. The notation $\mathrm{vec}(\bullet)$ is an isomorphism by which a matrix $X \in \mathbb{M}^{m\times n}$ with the ith row x^i $(i = 1, \ldots, m)$ is mapped to $[x^1 \cdots x^m]^{\mathrm{T}} \in \mathbb{R}^{mn}$ [12].

The new special basis is referred to as a Walsh basis in this brief. The basis is made by the Walsh functions and a family of the Hadamard–Walsh transform representation, which were popular [13], [14] in signal or image processing blocks but not today, because more precise and heavy transforms are implementable in the off-line world at least. On the other hand, the Walsh basis has not been applied for modeling and control of dynamical systems by the conventional visual feedbacks without geometric features.

In our visual feedback, since the Walsh basis elements are in the order of the space resolution (spatial resolution), strictly speaking, in the order of the horizontal first and the vertical second sequence (the number of the switch between the white and the black in the horizontal or vertical scanning), the Walsh basis is systematically truncated without geometric features. In addition, even though the original Walsh–Hadamard transform size $m \times n$ (the number of the basis elements mn) is not free as defined earlier, based on the projection theorem [8], the Walsh basis is freely truncated so that fruitful control theories are applicable.

The major difference between the Walsh basis in this brief and the alternative basis is the experimental cost. Unlike the Walsh basis, the alternative basis is generated by acquiring the experimental movies with a lot of specific information about the plant block. In return, the number of the alternative basis elements (the output dimension) can be smaller than that of the Walsh basis elements. In a word, the online experimental cost is reduced by the alternative basis, whereas the off-line experimental cost is reduced by the Walsh basis. Also, unlike the alternative basis, the Walsh basis is applicable to model free control (e.g., the PID control) skipping any off-line procedure. The range of the basis design will be increased by this brief.

III. Application to Nonplanar Sloshing

A. Experimental Setup

Sloshing [15], [16] is an important dynamical system in control systems technology [17]–[19]. Especially for nonplanar sloshing [16], [20], [21], the whole shape of the free surface is difficult to be measured by a few level sensors. As nonplanar sloshing is called nonlinear sloshing [15], [22], apart from numerical or experimental validations [23], the closed-loop stability has been difficult to be guaranteed. In a related paper [18], the whole shape of nonplanar surface is defined as a geometric feature and extracted in the image processing block. Since the whole shape of nonplanar surface is given in the control block, a model-based feedback is achieved as long as a certain input–output linearity exists on polynomial space. However, the design procedures of the image processing block are not systematic due to the geometric feature. In this brief, unlike in the related paper, even when the whole shape of nonplanar surface is not given in the control block in the presence of occlusion, a model-based feedback is achieved without geometric features. The control block and the image processing block are unified, and both design procedures are systematic.

Fig. 4 shows the system configuration. The calculation block is constructed with a real-timed control PC (Linux, 2.66 [GHz], 32 [b]) with the sampling rate $1/T_{\mathrm{sam}} = 15$ [Hz], a D/A board (12 [b]), and an image capture board (RGB, $8 \times 8 \times 8$ [b]). The actuation block is constructed with a dc motor (110 [W], 0.183 [Nm/A]), a reduction gear (31.155 [Nm/Nm]), and a current servo amplifier (1.5 [A/V]). The input voltage has the saturation (± 5 [V]). The plant block is constructed with a tank (glass, width 450 [mm] $\times$ long 180 [mm] $\times$ height 300 [mm]), water (blue, 0.998 [g/ml (20°)], 8.10 [L], depth 120 [mm]), liquid paraffin (colorless, 0.868 [g/ml (20°)], 12.15 [L], depth 180 [mm]), and a stage

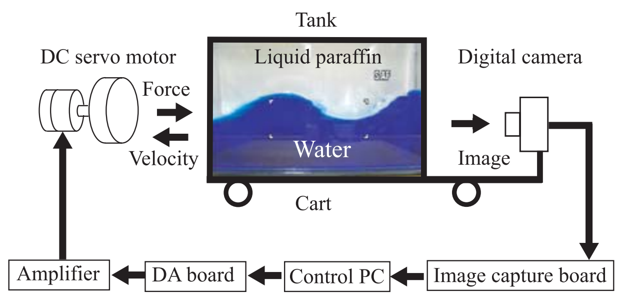

Fig. 4. System configuration.

cart. The driving torque of the dc motor is converted to the horizontal driving force for the tank dynamics (the camera and object dynamics) via a rack and a pinion (radius 100 [mm]). The static gain from the input voltage to the driving force is 172.04 [N/V].

The detection block is constructed with a digital camera under a room light (250 [lux]). The camera is allocated to detect the front view of the tank. Due to the computational limitation, every raw camera image (640 $\times$ 480 [pixel]) is reduced to a new camera image in $\mathbb{M}^{50\times 50}$ for evaluation only and not for control. More precisely, in a geometrical central part (600 $\times$ 450 [pixel]) of the raw camera image, the mean luminance of the several raw pixels (12 $\times$ 9 [pixel]) is replaced by a luminance of a new and larger pixel. This camera image reduction is not a part of the image processing block in the sense that the reduction is equivalent to a replacement of the original plant block with the raw camera by a virtual plant block with the new camera. The Walsh basis is generated in case of $N = 2^1 \times 2^1 = 4$ as a low-resolution case and $N = 2^3 \times 2^3 = 64$ as a high-resolution case so that our feedbacks are implementable. Accordingly, the raw camera image is reduced to another new camera image in $\mathbb{M}^{8\times 8}$ for control. In case of $N = 2^4 \times 2^4$, our feedbacks are not implementable due to the computational limitation.

B. System Identification

The identification input component is a chirplike signal

$$U(k) = \left(A_1 + \frac{A_2 - A_1}{M}t\right) \times \sin\left(2\pi\left(f_1 + \frac{f_2 - f_1}{M}t\right)t\right) E_1^W \tag{3}$$

with $t = T_{\mathrm{sam}}k$. The initial condition is the steady horizontal surface whose image Y_0 is similar to the element E_2^W in Fig. 3. Every output $Y(k)$ is the difference between the reduced new camera image for control and the steady horizontal surface image Y_0. The Walsh basis gives the output components $y_\ell(k) = \langle Y(k), E_\ell^W\rangle$ by (2). Note that our visual feedback
AQ:2 is geometric feature less but not feature less. Indeed, y_ℓ is a nongeometric feature.

Figs. 5–9 show the actual output components (the black dots) in case of $A_1 = 1.0$ [V], $A_2 = 2.0$ [V], $f_1 = 0.18$ [Hz], $f_2 = 0.90$ [Hz], and $M = 60$ [s]. The output component of the basis element E_1^W has an offset. This nonlinearity is due to the room light perturbation but the magnitude is not large relatively. The output components of the basis elements

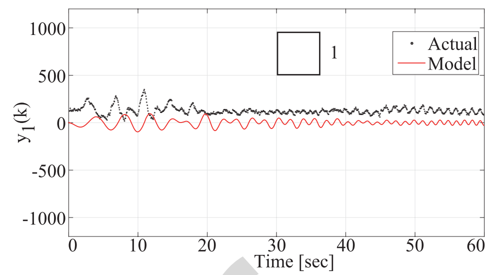

Fig. 5. Output components (E_1^W).

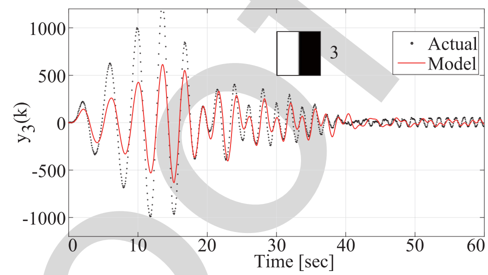

Fig. 6. Output components (E_3^W).

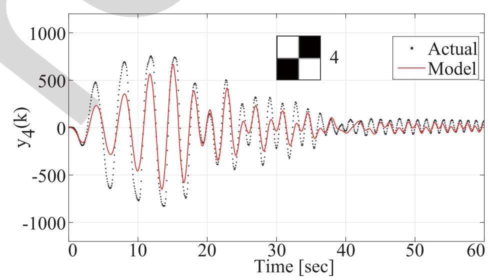

Fig. 7. Output components (E_4^W).

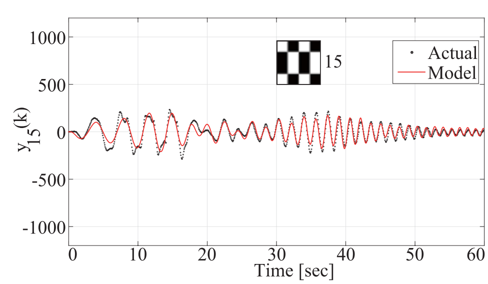

Fig. 8. Output components (E_{15}^W).

E_3^W and E_4^W are large at planar sloshing around $t = 10.0$ [s] and that of the basis elements E_{15}^W is also large at nonplanar sloshing around $t = 35.0$ [s]. On the other hand, the output component of the basis element E_{56}^W is always small relatively.

Fig. 10 shows the Bode plots. This is the result of the system identification (N4SID) to calculate the representation matrices of the mapping $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ of the controllable and observable system (1) with a 1-input component of $U(k) \in \mathbb{M}^{1\times 1}$, a 12-state component of $X(k) \in \mathbb{M}^{3\times 4}$, and a 64-output component of $Y(k) \in \mathbb{M}^{8\times 8}$. The size of the state matrix $X(k)$ is based on the representation size of $\mathcal{A}$. Note that the plant block in

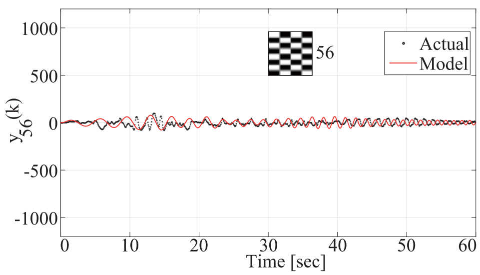

Fig. 9. Output components (E_{56}^W).

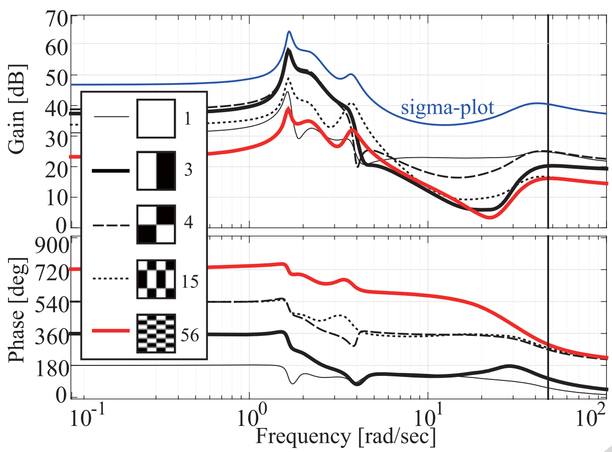

Fig. 10. Bode plot (identification).

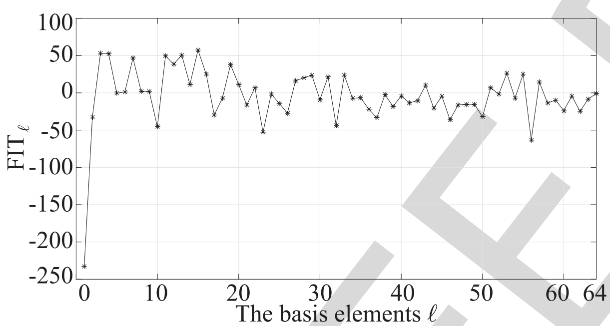

Fig. 11. Fit ratio (model validation).

case of $N = 4$ is of the form (1) with a 4-output component of $Y(k) \in \mathbb{M}^{2\times 2}$, which are the first four of the 64-output component of $Y(k) \in \mathbb{M}^{8\times 8}$ in case of $N = 64$. Every gain plot has the first peak at $\omega = 2\pi 0.285$ [rad/s]. Especially, the gain of the basis elements E_3^W and E_4^W is larger than the others around the peak. The gain plot of the basis element E_{15}^W has the second peak at $\omega = 2\pi 0.567$ [rad/s] unlike those of E_3^W and E_4^W. The first and the second peaks correspond to planar and nonplanar sloshing, respectively [9]. There are no additional peaks even in the (maximum) sigma plot taking all of the 64-output component. The gain plots of the basis elements E_ℓ^W ($\ell > 40$) are sufficiently small.

Fig. 11 shows the fit ratio [24]

$$\mathrm{FIT}_\ell := \left(1 - \sqrt{\frac{\Sigma\ \tilde{y}_\ell(k)^2}{\Sigma\,(y_\ell(k) - E[y_\ell(k)])^2}}\right) \times 100$$

where $\tilde{y}_\ell(k)$ is the difference between the actual output component $y_\ell(k)$ (the black dots) and the model output component (the red lines) in Figs. 5–9 by the same input. The notation $E[\bullet]$ denotes the expectation. The best fit ratio is achieved by the basis element E_{15}^W corresponding to nonplanar sloshing. The second and the third best fit ratios are achieved by the basis elements E_3^W and E_4^W corresponding to planar sloshing. These results imply that an input–output linearity exists on the matrix space. On the other hand, the worst and the secondary worst fit ratios are achieved by the basis elements E_1^W and E_{56}^W, respectively. This implies the existence of the uncertainty whose output is the state disturbance $V(k)$ in the input-state equation (1). However, both gains of the basis elements E_1^W and E_{56}^W in Fig. 10 are relatively small.

C. Control Experimental Method

The LQG control is applied on the matrix space. Fig. 12 shows the block diagram. We can skip the off-line basis generation procedure as well as the online geometric feature extraction. This simplicity is a fruit of our visual feedback. The control objective is the asymptotic stabilization of the plant origin, that is, the steady horizontal surface, in the presence of occlusion. The initial surface condition at $t = t_0 > 0$ is prepared by applying the feedforward input (3) with $A_1 = A_2$ and $f_1 = f_2$ in the period $[0, t_0]$ to the steady horizontal surface at $t = 0$. Here, we set $(A_1, f_1) = (0.9, 0.285)$ for planar sloshing and $(A_1, f_1) = (1.50, 0.567)$ for nonplanar sloshing, and $t_0 = 15$ [s]. Just after the feedforward input ends, we start the LQG control minimizing the objective functions [8]

$$\Sigma_0^\infty \left(q_f \langle X(k), X(k)\rangle + r_f \langle U(k), U(k)\rangle\right)$$

for the LQ controller and

$$E[\mathrm{vec}(X(k) - \hat{X}(k))\mathrm{vec}(X(k) - \hat{X}(k))^{\mathrm{T}}]$$

for the Kalman filter with the estimated state $\hat{X}(k)$ against the zero-mean disturbances $V(k)$ and $W(k)$ such that

$$E[\mathrm{vec}(V(k))\mathrm{vec}(V(k))^{\mathrm{T}}] = q_e I_{m_x n_x}$$
$$E[\mathrm{vec}(W(k))\mathrm{vec}(W(k))^{\mathrm{T}}] = r_e I_N$$

in which $(q_f, r_f) = (0.008, 30.77)$ and $(q_e, r_e) = (0.001, 10)$ in case of $N = 4$ (4-output), and $(q_f, r_f) = (0.0142, 17.61)$ and $(q_e, r_e) = (0.001, 50)$ in case of $N = 64$ (64-output), respectively. These weights q_f, r_f, q_e, and r_e are searched so that the inputs at planar sloshing take the same value at $t = 15$ [s] between $N = 4$ and $N = 64$ for a fair comparison. First, in the absence of occlusion, the stabilization by the proposed control is discussed. Second, in the presence of occlusion which is a student's hand, the rejection and the attenuation of the whole occlusion effects are also discussed.

D. Control Experimental Results and Discussion

Fig. 13 shows the input component of $U(k)$ in case of $N = 4$, and Fig. 14 shows the corresponding output norm $\|Y(k)\|$ for nonplanar sloshing in the absence of occlusion. The dot (black) depicts the no control, and the cross (red) depicts the proposed control. The output norm $\|Y(k)\|$ grows until the initial time $t = 15$ [s] by the feedforward input and converges

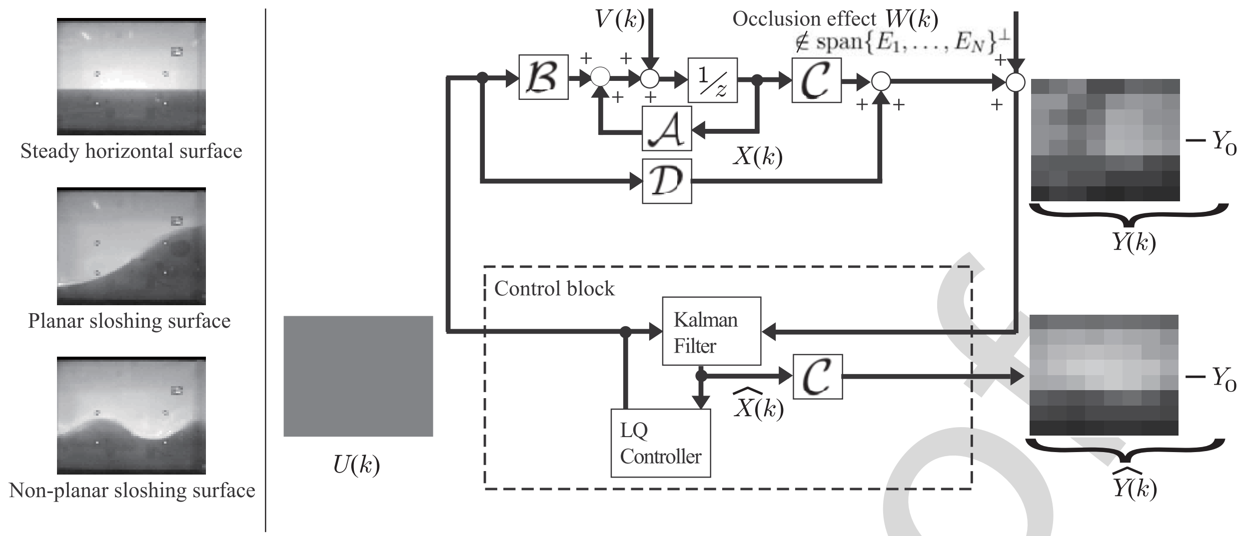

Fig. 12. Block diagram in which a part of the whole occlusion effects is rejected and the remaining part is attenuated.

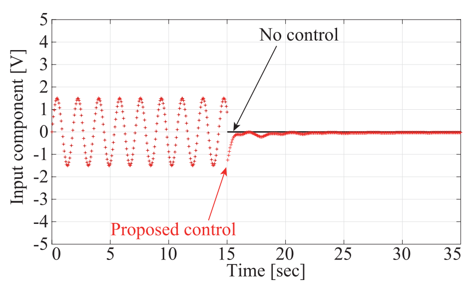

Fig. 13. Input component (without occlusion, $N = 4$).

TABLE I
MATRIX BASIS COMPARISON

Basis	Walsh $N{=}4$	Walsh $N{=}64$	POD $N{=}4$	POD $N{=}64$	No control
Cost (on-line)$^{\diamond}$	Low	High	Low	High	—
Cost (off-line)	**Low***	**Low***	High†	High†	—
Performance T_s	9.3	6.0	6.5	6.0	13.3

$^{\diamond}$The number of the basis elements (N)
*Two procedures (system identification, controller design)
†Three procedures (basis generation, system identification, controller design)

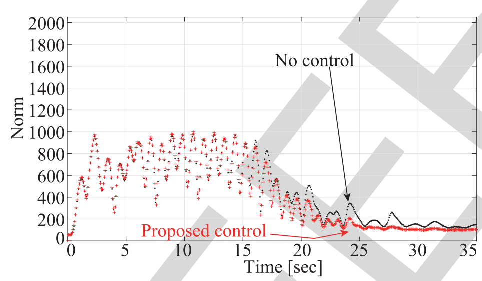

Fig. 14. Output norm (without occlusion, $N = 4$).

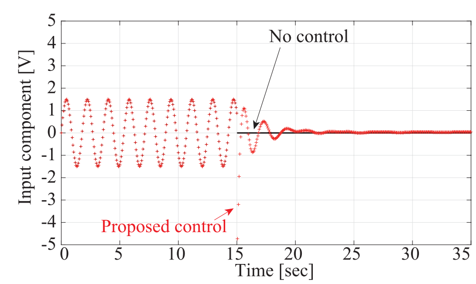

Fig. 15. Input component (without occlusion, $N = 64$).

to zero after $t = 15.0$ [s], as the input component converges to zero in the steady-state period. There is no input saturation. The convergence rate by the proposed control in case of $N = 4$ is slightly faster than that by the no control. The settling time by the proposed control in case of $N = 4$ is $T_s = 9.3$ [s] and that by the no control is $T_s = 13.3$ [s]. Here, the settling time T_s is a control performance introduced as the last time when the output norm is less than 20% of the maximum after we start the controls at $t = 15.0$ [s]. Note that Fig. 14 displays the high-resolution output for a fair comparison between $N = 4$ and $N = 64$.

Fig. 15 shows the input component of $U(k)$ in case of $N = 64$, and Fig. 16 shows the corresponding output norm $\|Y(k)\|$ for nonplanar sloshing in the absence of occlusion. There is no input saturation again. The settling time by the proposed control in case of $N = 64$ is $T_s = 6.0$ [s] and that by the no control is $T_s = 13.3$ [s] again. Especially, in the transient period $15 \leq t \leq 20$ [s], the convergence rate in case of $N = 64$ is much better than that in case of $N = 4$ successfully. In other words, the proposed control in case of $N = 4$ does not work well for nonplanar sloshing. This is because the space resolutions of the four basis elements $E_1^W, \ldots, E_4^W$ in Fig. 3 are lower than the others.

Table I summarizes the off-line and online experimental costs and the performance. The Walsh basis in case of $N = 64$ achieves the best performance. Here, $N = 64 (> 40)$ is very high so that the exchange of the Walsh basis for the alternative (POD) basis [9] can correspond to the change of basis and can

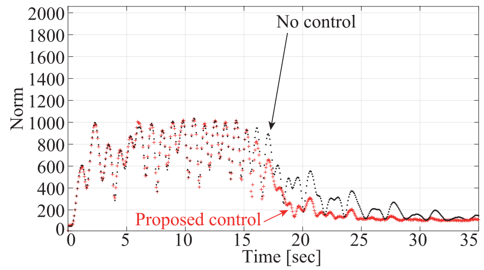

Fig. 16. Output norm (without occlusion, $N = 64$).

preserve the output norm

$$\|\mathcal{U} \circ Y(k)\| = \|Y(k)\|$$

for a unitary operator $\mathcal{U}$. This means that the alternative basis in case of $N = 64$ brings the same sigma plot in Fig. 10 and thus achieves the same performance $T_s = 6.0$ [s] experimentally in spite of the worse off-line experimental cost. On the other hand, the Walsh basis in case of $N = 4$ achieves the best experimental costs but has the worst performance than the alternative basis in case of $N = 4$. This is convincing, because the Walsh basis has no specific information about the plant block. In return, the Walsh basis can skip the off-line basis generation, and we do not have to handle the movies. The range of the basis design is increased successfully. The related paper [18] brings $T_s = 4.6-5.1$ [s] ($6.9-7.7$ [s] in the presence of occlusion) but is not fair here, because the design procedures of the image processing block are not systematic.

Fig. 17(a) shows the camera images for evaluation at the several discrete time by the proposed control in case of $N = 64$ in the absence of occlusion. The initial nonplanar surface tends to the plant origin. The validity is confirmed in the absence of occlusion.

On the other hand, Fig. 17(b) shows the camera images by the proposed control in case of $N = 64$ in the presence of occlusion. Fig. 17(c) shows the actual output $Y(k) + Y_0$, and Fig. 17(d) shows the estimated output $\hat{Y}(k) + Y_0$, which is calculated by the Kalman filter for the evaluation only and not for control. Here, the steady horizontal surface image Y_0 is added just for readability. Successfully, the initial nonplanar surface tends to the plant origin again. Also, by comparing Fig. 17(c) and Fig. 17(d), the occlusion effect in the output $Y(k)+Y_0$ is attenuated and almost disappeared in the estimated output $\hat{Y}(k)+Y_0$. Not only the plant origin but also the closed-loop origin is asymptotically stabilized.

Remark 3: Unlike conventional visual feedbacks, our visual feedback handles the occlusion in two steps without geometric features. In the first step, in the sense that the occlusion in Fig. 17(a) is projected to the occlusion effect that exists in $\mathbb{W} := \text{span}\{E_1^W, \ldots, E_{64}^W\}$, a part of the whole occlusion effects exists in $\mathbb{W}^{\perp}$ experimentally and is rejected. In the second step, the remaining part exists in $\mathbb{W}$ and is attenuated by the LQG control. The control performance in the presence of occlusion depends not only on the controller transfer function in the second step but also on the relation between the occlusion and the basis in the first step. There is no basis whose occlusion effect rejection performance is always better than the others for every possible occlusion.

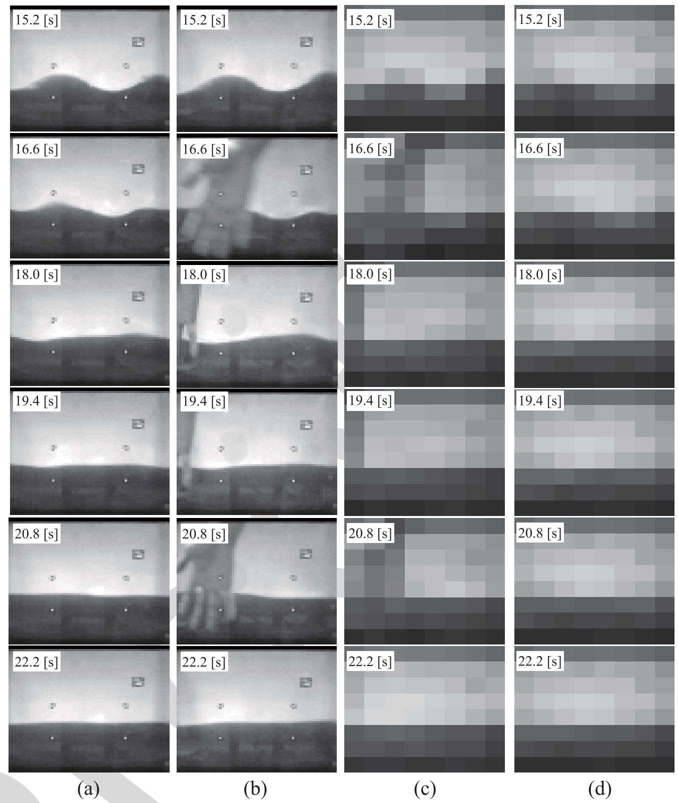

Fig. 17. Images and outputs. (a) Images (without occlusion). (b) Images (with occlusion). (c) $Y(k) + Y_0$ (with occlusion). (d) $\hat{Y}(k) + Y_0$ (with occlusion).

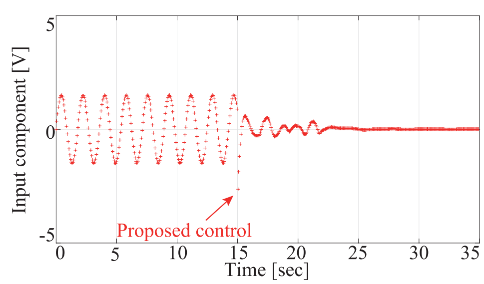

Fig. 18. Input component (with occlusion, $N = 64$).

Figs. 18 and 19 show the input component and the output norm in the presence of occlusion for nonplanar sloshing discussed in Fig. 17(a)–(c). These settling times are slightly larger than those without occlusion. Especially in the transient period $15 \leq t \leq 25$ [s], the existence of the occlusion is observed, but the input and output components tend to be zero in the steady-state period again. The validity is confirmed even in the presence of occlusion. As a demonstration, Fig. 20 shows the output norm $\|Y(k)\|$ against the input disturbance [the same chirplike input (3) for the system identification] instead of the output disturbance (the occlusion effect). Again, the proposed control is better than the no control.

Finally, let us discuss the robust stability analysis. This is also a demonstration that our visual feedback guarantees the closed-loop stability even in the presence of the (input multiplicative) uncertainty Δ and the occlusion effect W. Taking the extended block structure set: $\mathbf{\Delta} = \{\text{diag}(\Delta, \Delta_f) \mid \Delta \in \mathbb{C}, \Delta_f \in \mathbb{C}^{N\times N}\}$, it is known that the robust performance as

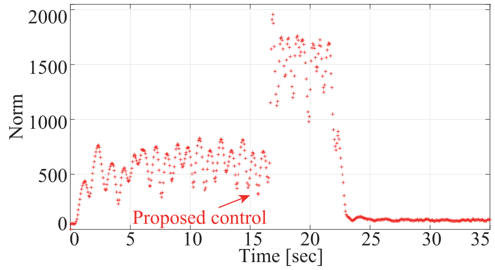

Fig. 19. Output norm (with occlusion, $N = 64$).

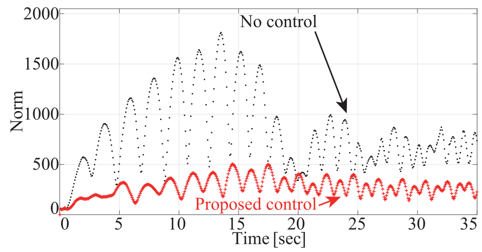

Fig. 20. Output norm (with input disturbance, $N = 64$).

well as the robust stability is evaluated by structured singular value (SSV): $\mu_{\mathbf{\Delta}}(G) = (\min\{\bar{\sigma}(\Delta) \mid \Delta \in \mathbf{\Delta},\ \det(I - G\Delta) = 0\})^{-1}$ with a standard setting [12]

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \underbrace{\begin{bmatrix} -KS_oP & -KS_o \\ S_oP & S_o \end{bmatrix}}_{G} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$S_o = (I + L_o)^{-1}, \quad L_o = PK$$

where P is the nominal plant block from the input U to the output Y in the absence of W, K is the control block from the disturbed output Y to the input U of the perturbed plant block $P(1 + \Delta)$, z_1 is the input U, z_2 is the disturbed output Y, $w_1 = \Delta z_1$, and w_2 is the output disturbance W. By the standard μ-analysis (MATLAB version 9.0) in the worst case of $N = 4$ and 64, in the presence of the uncertainty Δ with the observed maximum gain 2.087, one of the upper bounds of SSV is lower than 0.3612, which guarantees both of the robust stability for $|\Delta| \le 2.087 < 1/0.3612$ and the robust performance $\|F_u(G, \Delta)\|_\infty < 0.3612$. The notation $F_u(\bullet, \bullet)$ denotes the upper linear fractional transformation. Unlike in the LQG controller design procedure, the zero-mean assumption is not needed, and the output of the uncertainty is the input disturbance. The tools (N4SID, LQG, and μ) in this brief are examples, and various other tools on the coordinate space are applicable to other dynamical systems (e.g., other continuum systems) on the matrix space as long as the input–output linearity exists in the sense of the system (1).

IV. Conclusion

For a visual feedback without geometric features, this brief suggests to apply a new special basis made by the Walsh functions to reduce the off-line experimental cost. The validity is confirmed experimentally against occlusion in nonplanar sloshing whose dynamics is not negligible. The range of the basis design is increased. The next work is a systematic basis generation to improve the input–output linearity as well as the occlusion effect rejection performance assuming that some *a priori* information about the occlusion is available.

Acknowledgment

The authors would like to thank the reviewers for the valuable comments, and they would also like to thank K. Nagai, S. Morozumi, and K. Lee for their experimental helps.

References

[1] G. Silveira and E. Malis, “Direct visual servoing: Vision-based estimation and control using only nonmetric information,” *IEEE Trans. Robot.*, vol. 28, no. 4, pp. 974–980, Aug. 2012.
[2] [illegible]. Zang and K. Hashimoto, “GPU acceleration in a visual servo system,” *[illegible] Robot. Mech.*, vol. 24, no. 1, pp. 105–114, 2012.
[3] J. Zhao, B. Song, N. Xi, L. Sun, H. Chen, and Y. Jia, “Non-vector space approach for nanoscale motion control,” *Automatica*, vol. 50, no. 7, pp. 1835–1842, 2014.
[4] K. Deguchi and T. Noguchi, “Visual servoing using eigenspace method and dynamic calculation of interaction matrices,” in *Proc. IEEE ICPR*, Aug. 1996, pp. 302–306.
[5] S. K. Nayar, S. A. Nene, and H. Murase, “Subspace methods for robot vision,” *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 750–758, Oct. 1996.
[6] A. Censi, S. Han, S. B. Fuller, and R. M. Murray, “A bio-plausible design for visual attitude stabilization,” in *Proc. IEEE CDC*, Dec. 2009, pp. 3513–3520.
[7] S. Han, A. Censi, A. D. Straw, and R. M. Murray, “A bio-plausible design for visual pose stabilization,” in *Proc. IEEE/RSJ IROS*, Oct. 2010, pp. 5679–5686.
[8] D. G. Luenberger, *Optimization by Vector Space Methods*. New York, NY, USA: Wiley, 1960.
[9] S. Sakai and M. Ando, “On the visual systems & control on matrix space,” in *Proc. IEEE CDC*, Dec. 2014, pp. 2173–2178.
[10] K. Nagai and S. Sakai, “A visual feedback design on matrix space for a liquid sloshing experiment,” in *Proc. SICE*, Sep. 2013, pp. 2088–2093.
[11] J. L. Walsh, “A closed set of normal orthogonal functions,” *Amer. J. Math.*, vol. 45, no. 1, pp. 5–24, 1923.
[12] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
[13] K. G. Beauchamp, *Walsh Functions and Their Applications*. Norwell, MA, USA: Academic, 1975.
[14] W. K. Pratt, J. Kane, and H. C. Andrews, “Hadamard transform image coding,” *Proc. IEEE*, vol. 57, no. 1, pp. 58–68, Jan. 1969.
[15] R. A. Ibrahim, *Liquid Sloshing Dynamics: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[16] T. Ikeda, M. Takashima, and Y. Harata, “Autoparametric resonances of elastic structures coupled with two sloshing modes in a square liquid tank,” *ASME J. Comput. Nonlinear Dyn.*, vol. 8, no. 1, p. 011007, 2013.
[17] B. Robu, L. Baudouin, C. Prieur, and D. Arzelier, “Simultaneous H_∞ vibration control of fluid/plate system via reduced-order controller,” *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 3, pp. 700–711, May 2012.
[18] S. Sakai and M. Sato, “Visual systems & control on polynomial space and its application to sloshing problems,” *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 6, pp. 2176–2187, Nov. 2014.
[19] K. Yano and K. Terashima, “Robust liquid container transfer control for complete sloshing suppression,” *IEEE Trans. Control Syst. Technol.*, vol. 9, no. 3, pp. 483–493, May 2001.
[20] M. Grundelius and B. Bernhardsson, “Control of liquid slosh in an industrial packaging machine,” in *Proc. Int. Conf. Control Appl.*, Aug. 1999, pp. 1654–1659.
[21] M. Peric, T. Zorn, O. El Moctar, T. E. Schellin, and Y.-S. Kim, “Simulation of sloshing in LNG-tanks,” *ASME J. Offshore Mech. Arctirce Eng.*, vol. 131, no. 3, pp. 031101-1–031101-11, 2009.
[22] O. M. Faltinsen and A. N. Timokha, “On sloshing modes in a circular tank,” *J. Fluid Mech.*, vol. 695, pp. 467–477, Mar. 2012.
[23] Q. Zang, J. Huang, and Z. Liang, “Slosh suppression for infinite modes in a moving liquid container,” *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 1, pp. 217–225, Feb. 2015.
[24] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

AQ:3

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

PLEASE NOTE: We cannot accept new source files as corrections for your paper. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.

AQ:1 = Please provide the expansion for the acronyms “ARX and N4SID.”
AQ:2 = Please confirm whether the retention of the sentence “Note that our visual feedback is geometric feature less but not feature less” is correct.
AQ:3 = Please confirm the journal title, volume no., issue no., page range, and year for ref. [2].

Visual Feedback Without Geometric Features Against Occlusion: A Walsh Basis

Satoru Sakai, *Member, IEEE*, Masayuki Ando, and Shunsuke Kobashi

***Abstract*— For a visual feedback without geometric features, this brief suggests to apply a basis made by the Walsh functions in order to reduce the off-line experimental cost. Depending on the resolution, the feedback is implementable and achieves the closed-loop stability of dynamical systems as long as the input–output linearity on matrix space exists. Remarkably, a part of the whole occlusion effects is rejected, and the remaining part is attenuated. The validity is confirmed by the experimental feedback for nonplanar sloshing.**

***Index Terms*— Dynamical systems, occlusion, stability, visual feedback.**

I. Introduction

IN MANY conventional visual feedbacks, there exists a series interconnection between the control block and the image processing block in Fig. 1. In the image processing block, the geometric features (e.g., a dot position and a line angle) are defined and extracted from the camera image on line. Via the series interconnection, a lot of information is lost in the image processing block, but the design procedures of the control block can be systematic when fruitful control theories are applicable. On the other hand, the design procedures of the image processing block are not or less systematic, especially in the presence of occlusion (visual obstacles between the camera and the object), because the way to define and extract geometric features strongly depends on the plant block, the control objective, and so on.

To solve this problem, not many but several visual feedbacks without or with less geometric features are discussed by different approaches, such as the homography-based approach [1], [2] and the Hausdorff distance-based approach [3]. The similar motivation is traced back to the subspace approaches [4], [5]. Most of them could work locally at least for static systems that are acceptable when the camera or object dynamics (e.g., the camera-link flexibility) are negligible. On the other hand, the closed-loop stability of dynamical systems is not guaranteed and can be lost even in the absence of occlusion. Exceptionally, a visual feedback [6], [7] locally guarantees the closed-loop stability of a special nonlinear dynamical system assuming the absence of occlusion.

In this brief, in the presence of occlusion, a visual feedback without geometric features is given as a new application for linear dynamical systems. The closed-loop stability is guaranteed by our simple idea beginning with a change of the mapping domain and codomain (the input and output spaces) of the plant block. In many conventional visual feedbacks, geometric features are defined in a coordinate space $\mathbb{R}^n$ (e.g., the camera image plane $\mathbb{R}^2$), which can be eventually the output space of the plant block. In our visual feedback in Fig. 2, geometric features are not defined, and a matrix space $\mathbb{M}^{m\times n}$ is the output space of the plant block. Since any coordinate space is isomorphic to a matrix space, the design procedures of our control block on matrix space can be systematic when fruitful control theories are applicable again.

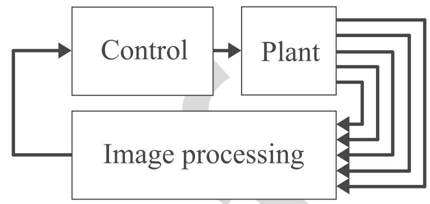

Fig. 1. Block diagram with the image processing for geometric features.

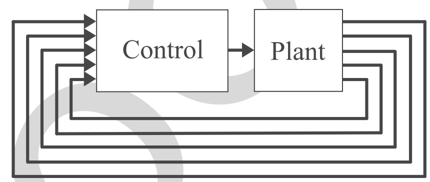

Fig. 2. Block diagram without the image processing for geometric features.

However, due to the computational limitation at least, such theories are not always applicable as they are. In our visual feedback, from the perspective of the Hilbert space [8], we can design a basis in the output space $\mathbb{M}^{m\times n}$ so that the control theories are applicable under the computational limitation. Indeed, in the absence of occlusion, our pilot study [9] performs an off-line basis generation procedure before the system identification procedure.

In the presence of occlusion, this brief suggests to apply a new special basis by which any off-line basis generation procedure is not needed. This means a cut of the experimental cost, because the experimental movies for the off-line basis generation procedure are nothing but big data for control. The new special basis is made by the Walsh functions, which have not been applied for modeling and control of dynamical systems by the conventional visual feedbacks without geometric features.

The rest of this brief is organized as follows. In Section II, dynamical systems on matrix space are introduced, and the new special basis is suggested for our visual feedback.

Manuscript received January 24, 2017; revised October 2, 2017; accepted November 30, 2017. Manuscript received in final form December 1, 2017. This work was supported by JSPS KAKENHI under Grant 17K06227. Recommended by Associate Editor A. Behal. *(Corresponding author: Satoru Sakai.)*

The authors are with the Department of Mechanical Engineering, Shinshu University, Matsumoto 4-17-1, Japan (e-mail: satorusakai@gakushikai.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCST.2017.2780176

The new special basis does not need any off-line basis generation procedure but can be systematically truncated without geometric features so that fruitful control theories are applicable under the computational limitation. In Section III, the proposed control is applied to nonplanar sloshing whose dynamics is not negligible. The validity is confirmed experimentally in the presence of occlusion. Finally, this brief is concluded in Section IV.

II. Dynamical Systems on Matrix Space

Let us consider a finite-dimensional space denoted by $\mathbb{M}^{m\times n}$ of a time-varying matrix $M(k) \in \mathbb{M}^{m\times n}$ at the discrete-time instant $k \in \mathbb{Z}_+ := \{0\}\cup\mathbb{N}$. The matrix space $\mathbb{M}^{m\times n}$ is a Hilbert space with the inner product

$$\langle M(k), N(k)\rangle = \mathrm{tr}(M(k)^{\mathrm{T}} N(k)) \in \mathbb{R}$$

for all matrices $M(k)$ and $N(k) \in \mathbb{M}^{m\times n}$. $M(k) \perp N(k)$ implies $\langle M(k), N(k)\rangle = 0$, and the inner product introduces the norm $\|M(k)\| = (\langle M(k), M(k)\rangle)^{1/2} \geq 0$. The notation $\mathrm{tr}(\bullet)$ denotes the trace of a matrix. Consider a finite-dimensional linear time-invariant (LTI) system described by linear mappings between matrix spaces [9]

$$\begin{cases} X(k+1) = \mathcal{A}\circ X(k) + \mathcal{B}\circ U(k) + V(k) \\ Y(k) = \mathcal{C}\circ X(k) + \mathcal{D}\circ U(k) + W(k) \end{cases} \tag{1}$$

where the state $X(k) \in \mathbb{M}^{m_x\times n_x}$ and the state disturbance $V(k) \in \mathbb{M}^{m_x\times n_x}$ are the $m_x \times n_x$ matrices, the input $U(k) \in \mathbb{M}^{1\times 1}$ is the 1×1 matrix, and the output $Y(k) \in \mathbb{M}^{m_y\times n_y}$ and the output disturbance $W(k) \in \mathbb{M}^{m_y\times n_y}$ are the $m_y \times n_y$ matrices. The notation $\circ$ denotes the operation of the linear mappings $\mathcal{A}, \mathcal{B}, \mathcal{C}$, and $\mathcal{D}$.

Remark 1: Since every mapping cannot be defined until the domain and the codomain are defined, every system depends on the choice of the input and output spaces. In this sense, the proposed system (1) choosing the matrix spaces and the well-known LTI system choosing the coordinates spaces are different mathematical objects even if the linear mappings $\mathcal{A}, \mathcal{B}, \mathcal{C}$, and $\mathcal{D}$ of both systems have the same matrix representations. On the other hand, since there is an isomorphism from a matrix space $\mathbb{M}^{m\times n}$ to a coordinate space $\mathbb{R}^{mn}$ [8], fruitful control theories [e.g., ARX, N4SID, linear quadratic and Gaussian (LQG), and μ] are applicable to both systems.

Remark 2: The output Y corresponding to the camera image is visible, but the input U and the state X are invisible as they are the disturbances V and W. Of course, the input U is not unknown and visualizable, but the state X is not always visualizable even in the absence of the disturbances V and W.

It is never our contribution to see the camera image as a matrix and is popular in the image processing blocks that are regarded as static open systems. In our visual feedback, from the viewpoint of dynamical closed-loop systems including the plant block, not only the camera image corresponding to the output of the plant block, but also the input and state are matrices. The inner product (or the passivity) of the output Y and the input U can be taken when they belong to the same subspace. In mathematics, roughly speaking, a matrix space is almost the same as a coordinate space, which is familiar. However, in engineering, as long as the control objective is defined in the camera image, the matrix space is more suitable to design the basis.

Since a matrix space $\mathbb{M}^{m\times n}$ has a normalized orthogonal basis $E_1, \ldots, E_{mn}$ [8]

$$\langle E_{\ell_i}, E_{\ell_j}\rangle = \begin{cases} 0 \ (\ell_i \neq \ell_j) \\ 1 \ (\ell_i = \ell_j), \end{cases} \qquad \ell_i, \ell_j = 1, \ldots, mn$$

every time-varying matrix

$$M(k) = \sum_{\ell=1}^{mn} \langle M(k), E_\ell\rangle E_\ell \in \mathbb{M}^{m\times n}, \quad \ell = 1, \ldots, mn$$

has a representation $[m_1(k), m_2(k), \ldots, m_{mn}(k)]^{\mathrm{T}}$ whose component is of the form

$$m_\ell(k) := \langle M(k), E_\ell\rangle. \tag{2}$$

Here, the most popular basis in the output space is the standard basis (the pixel-by-pixel basis)

$$E_1^S = \begin{bmatrix} 1\cdots 0 \\ \vdots \quad \vdots \\ 0\cdots 0 \end{bmatrix}$$

$$E_2^S = \begin{bmatrix} 0 \ 1\cdots 0 \\ \vdots \ \vdots \quad \vdots \\ 0 \ 0\cdots 0 \end{bmatrix}, \ldots, E_{mn}^S = \begin{bmatrix} 0\cdots 0 \\ \vdots \quad \vdots \\ 0\cdots 1 \end{bmatrix}$$

by which any off-line basis generation procedure is not needed. The standard basis could work locally at least for static systems as the pixel-by-pixel feedback. However, the standard basis can cause several problems for dynamical systems. One of them is from the computational limitation, because the number of the standard basis elements is nothing but the number of the pixels mn, which is usually quite large [10]. Indeed, a more than 1×10^6 pixels feedback is implemented on a better hardware [2]. Nevertheless, the standard basis cannot be truncated systematically without geometric features. For example, for a certain plant block with a control objective, even if we know that the (1,2)-pixel of the camera image is not important, the truncation of E_2^S is not accepted, because such truncation is nothing but the geometric feature extraction depending on the plant block or the control objectives.

To solve the standard basis problem, under the computational limitation, our pilot study [9] discusses an alternative basis, which is systematically truncated without geometric features. However, the alternative basis needs an off-line basis generation procedure before the system identification procedure. This means an increase of the experimental cost, since the alternative basis cannot be generated without acquiring the experimental movies.

One may think that the experimental cost in the off-line basis generation procedure is not an issue, since the acquired movies for the off-line basis generation procedure can be reused for the system identification procedure. This is not true. The acquired movies for the off-line basis generation procedure are nothing but big data for control (e.g., the raw movies) and are much bigger than the outputs for the system

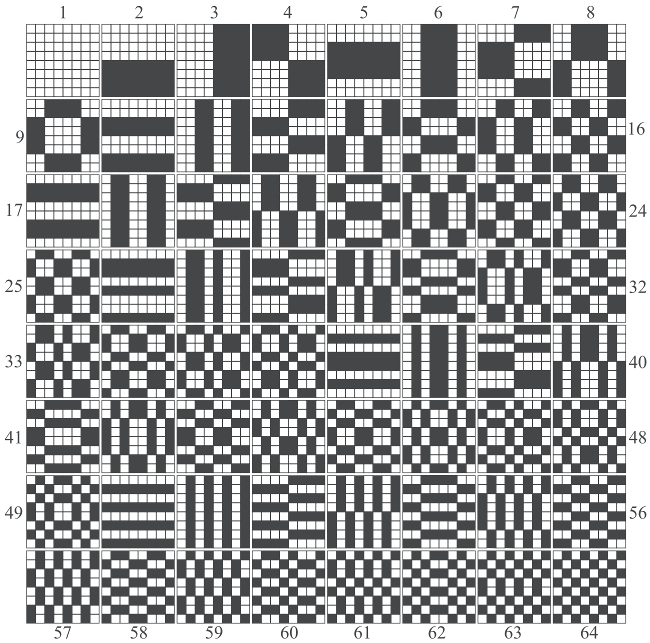

Fig. 3. Walsh basis in the order of the horizontal first and the vertical second sequence (white $= +1/64$ and black $= -1/64$).

identification procedure in which the number of the basis elements (the output dimension) is already fixed.

To solve the alternative basis problem as well as the standard basis problem, this brief suggests to apply a new special basis, which can be systematically truncated without geometric features but does not need any off-line basis generation procedure. Fig. 3 shows an example of the new special basis

$$E_\ell^W = \mathrm{vec}^{-1}\left(\left[W\left(\ell-1,\frac{0}{mn}\right)\cdots W\left(\ell-1,\frac{mn-1}{mn}\right)\right]^{\mathrm{T}}\right)$$

with

$$\underbrace{W(\ell-1,x) = (-1)^{\Sigma_{j=1}^{\infty} x_j(\ell-1)_{1-j}}}_{\text{the Walsh function [11]}},\quad \ell = 1,\ldots,mn$$

whose $x_j \in \{0,1\}$ and $(\ell-1)_{1-j} \in \{0,1\}$ are determined by the dyadic expansion of the normalized space parameter $x = \Sigma_{j=1}^{\infty} x_j \cdot 2^{-j} \in [0,1)$ and that of the number $(\ell-1) = \Sigma_{j=0}^{K}(\ell-1)_{-j} \cdot 2^j \in \mathbb{Z}_+$ $(K \in \mathbb{Z}_+, (\ell-1) \in [2^K, 2^{K+1}))$. Here, the number $N := mn$ is constrained to be $m = n = 2^L$ $(\forall L \in \mathbb{Z}_+)$. The notation vec$(\bullet)$ is an isomorphism by which a matrix $X \in \mathbb{M}^{m\times n}$ with the ith row x^i $(i = 1,\ldots,m)$ is mapped to $[x^1 \cdots x^m]^{\mathrm{T}} \in \mathbb{R}^{mn}$ [12].

The new special basis is referred to as a Walsh basis in this brief. The basis is made by the Walsh functions and a family of the Hadamard–Walsh transform representation, which were popular [13], [14] in signal or image processing blocks but not today, because more precise and heavy transforms are implementable in the off-line world at least. On the other hand, the Walsh basis has not been applied for modeling and control of dynamical systems by the conventional visual feedbacks without geometric features.

In our visual feedback, since the Walsh basis elements are in the order of the space resolution (spatial resolution), strictly speaking, in the order of the horizontal first and the vertical second sequence (the number of the switch between the white and the black in the horizontal or vertical scanning), the Walsh basis is systematically truncated without geometric features. In addition, even though the original Walsh–Hadamard transform size $m \times n$ (the number of the basis elements mn) is not free as defined earlier, based on the projection theorem [8], the Walsh basis is freely truncated so that fruitful control theories are applicable.

The major difference between the Walsh basis in this brief and the alternative basis is the experimental cost. Unlike the Walsh basis, the alternative basis is generated by acquiring the experimental movies with a lot of specific information about the plant block. In return, the number of the alternative basis elements (the output dimension) can be smaller than that of the Walsh basis elements. In a word, the online experimental cost is reduced by the alternative basis, whereas the off-line experimental cost is reduced by the Walsh basis. Also, unlike the alternative basis, the Walsh basis is applicable to model free control (e.g., the PID control) skipping any off-line procedure. The range of the basis design will be increased by this brief.

III. Application to Nonplanar Sloshing

A. Experimental Setup

Sloshing [15], [16] is an important dynamical system in control systems technology [17]–[19]. Especially for nonplanar sloshing [16], [20], [21], the whole shape of the free surface is difficult to be measured by a few level sensors. As nonplanar sloshing is called nonlinear sloshing [15], [22], apart from numerical or experimental validations [23], the closed-loop stability has been difficult to be guaranteed. In a related paper [18], the whole shape of nonplanar surface is defined as a geometric feature and extracted in the image processing block. Since the whole shape of nonplanar surface is given in the control block, a model-based feedback is achieved as long as a certain input–output linearity exists on polynomial space. However, the design procedures of the image processing block are not systematic due to the geometric feature. In this brief, unlike in the related paper, even when the whole shape of nonplanar surface is not given in the control block in the presence of occlusion, a model-based feedback is achieved without geometric features. The control block and the image processing block are unified, and both design procedures are systematic.

Fig. 4 shows the system configuration. The calculation block is constructed with a real-timed control PC (Linux, 2.66 [GHz], 32 [b]) with the sampling rate $1/T_{\mathrm{sam}} = 15$ [Hz], a D/A board (12 [b]), and an image capture board (RGB, $8 \times 8 \times 8$ [b]). The actuation block is constructed with a dc motor (110 [W], 0.183 [Nm/A]), a reduction gear (31.155 [Nm/Nm]), and a current servo amplifier (1.5 [A/V]). The input voltage has the saturation (± 5 [V]). The plant block is constructed with a tank (glass, width 450 [mm] $\times$ long 180 [mm] $\times$ height 300 [mm]), water (blue, 0.998 [g/ml (20°)], 8.10 [L], depth 120 [mm]), liquid paraffin (colorless, 0.868 [g/ml (20°)], 12.15 [L], depth 180 [mm]), and a stage

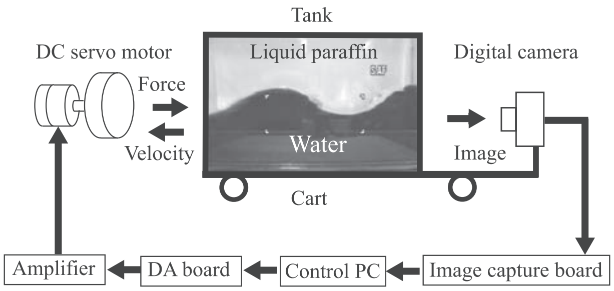

Fig. 4. System configuration.

cart. The driving torque of the dc motor is converted to the horizontal driving force for the tank dynamics (the camera and object dynamics) via a rack and a pinion (radius 100 [mm]). The static gain from the input voltage to the driving force is 172.04 [N/V].

The detection block is constructed with a digital camera under a room light (250 [lux]). The camera is allocated to detect the front view of the tank. Due to the computational limitation, every raw camera image (640 $\times$ 480 [pixel]) is reduced to a new camera image in $\mathbb{M}^{50\times 50}$ for evaluation only and not for control. More precisely, in a geometrical central part (600 $\times$ 450 [pixel]) of the raw camera image, the mean luminance of the several raw pixels (12 $\times$ 9 [pixel]) is replaced by a luminance of a new and larger pixel. This camera image reduction is not a part of the image processing block in the sense that the reduction is equivalent to a replacement of the original plant block with the raw camera by a virtual plant block with the new camera. The Walsh basis is generated in case of $N = 2^1 \times 2^1 = 4$ as a low-resolution case and $N = 2^3 \times 2^3 = 64$ as a high-resolution case so that our feedbacks are implementable. Accordingly, the raw camera image is reduced to another new camera image in $\mathbb{M}^{8\times 8}$ for control. In case of $N = 2^4 \times 2^4$, our feedbacks are not implementable due to the computational limitation.

B. System Identification

The identification input component is a chirplike signal

$$U(k) = \left(A_1 + \frac{A_2 - A_1}{M}t\right) \times \sin\left(2\pi\left(f_1 + \frac{f_2 - f_1}{M}t\right)t\right) E_1^W \tag{3}$$

with $t = T_{\mathrm{sam}}k$. The initial condition is the steady horizontal surface whose image Y_0 is similar to the element E_2^W in Fig. 3. Every output $Y(k)$ is the difference between the reduced new camera image for control and the steady horizontal surface image Y_0. The Walsh basis gives the output components $y_\ell(k) = \langle Y(k), E_\ell^W \rangle$ by (2). Note that our visual feedback is geometric feature less but not feature less. Indeed, y_ℓ is a nongeometric feature.

Figs. 5–9 show the actual output components (the black dots) in case of $A_1 = 1.0$ [V], $A_2 = 2.0$ [V], $f_1 = 0.18$ [Hz], $f_2 = 0.90$ [Hz], and $M = 60$ [s]. The output component of the basis element E_1^W has an offset. This nonlinearity is due to the room light perturbation but the magnitude is not large relatively. The output components of the basis elements

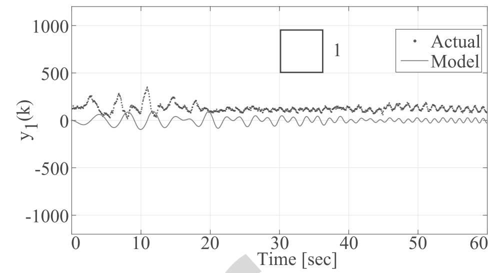

Fig. 5. Output components (E_1^W).

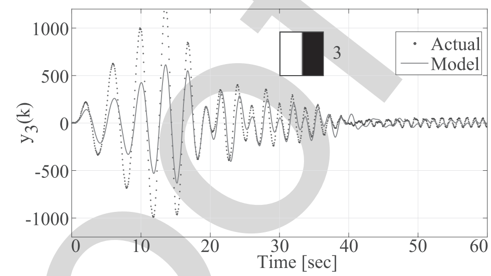

Fig. 6. Output components (E_3^W).

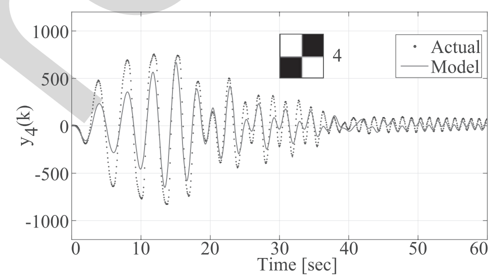

Fig. 7. Output components (E_4^W).

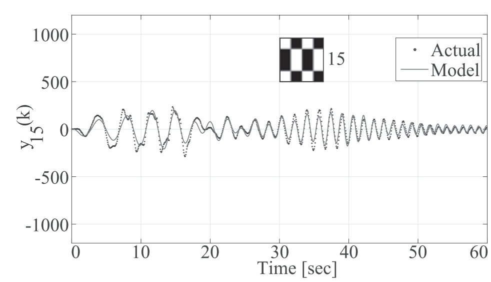

Fig. 8. Output components (E_{15}^W).

E_3^W and E_4^W are large at planar sloshing around $t = 10.0$ [s] and that of the basis elements E_{15}^W is also large at nonplanar sloshing around $t = 35.0$ [s]. On the other hand, the output component of the basis element E_{56}^W is always small relatively.

Fig. 10 shows the Bode plots. This is the result of the system identification (N4SID) to calculate the representation matrices of the mapping $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ of the controllable and observable system (1) with a 1-input component of $U(k) \in \mathbb{M}^{1\times 1}$, a 12-state component of $X(k) \in \mathbb{M}^{3\times 4}$, and a 64-output component of $Y(k) \in \mathbb{M}^{8\times 8}$. The size of the state matrix $X(k)$ is based on the representation size of $\mathcal{A}$. Note that the plant block in

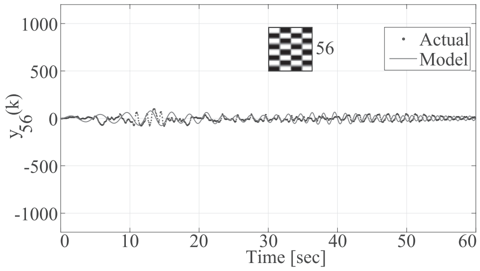

Fig. 9. Output components (E_{56}^W).

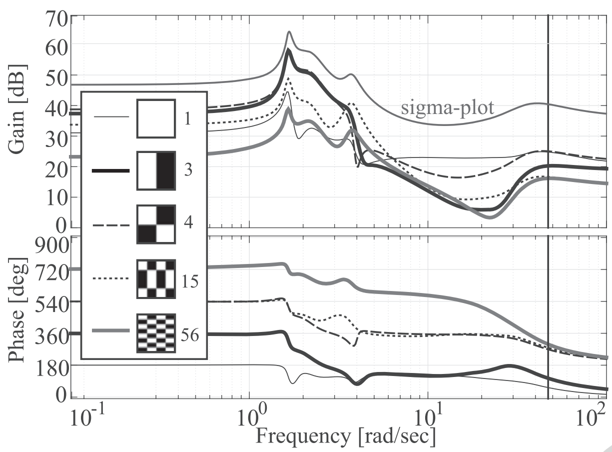

Fig. 10. Bode plot (identification).

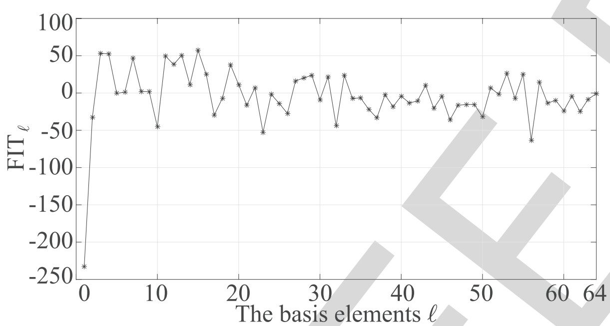

Fig. 11. Fit ratio (model validation).

case of $N = 4$ is of the form (1) with a 4-output component of $Y(k) \in \mathbb{M}^{2\times 2}$, which are the first four of the 64-output component of $Y(k) \in \mathbb{M}^{8\times 8}$ in case of $N = 64$. Every gain plot has the first peak at $\omega = 2\pi 0.285$ [rad/s]. Especially, the gain of the basis elements E_3^W and E_4^W is larger than the others around the peak. The gain plot of the basis element E_{15}^W has the second peak at $\omega = 2\pi 0.567$ [rad/s] unlike those of E_3^W and E_4^W. The first and the second peaks correspond to planar and nonplanar sloshing, respectively [9]. There are no additional peaks even in the (maximum) sigma plot taking all of the 64-output component. The gain plots of the basis elements E_ℓ^W ($\ell > 40$) are sufficiently small.

Fig. 11 shows the fit ratio [24]

$$\mathrm{FIT}_\ell := \left(1 - \sqrt{\frac{\Sigma\ \tilde{y}_\ell(k)^2}{\Sigma(y_\ell(k) - E[y_\ell(k)])^2}}\right) \times 100$$

where $\tilde{y}_\ell(k)$ is the difference between the actual output component $y_\ell(k)$ (the black dots) and the model output component (the red lines) in Figs. 5–9 by the same input. The notation $E[\bullet]$ denotes the expectation. The best fit ratio is achieved by the basis element E_{15}^W corresponding to nonplanar sloshing. The second and the third best fit ratios are achieved by the basis elements E_3^W and E_4^W corresponding to planar sloshing. These results imply that an input–output linearity exists on the matrix space. On the other hand, the worst and the secondary worst fit ratios are achieved by the basis elements E_1^W and E_{56}^W, respectively. This implies the existence of the uncertainty whose output is the state disturbance $V(k)$ in the input-state equation (1). However, both gains of the basis elements E_1^W and E_{56}^W in Fig. 10 are relatively small.

C. Control Experimental Method

The LQG control is applied on the matrix space. Fig. 12 shows the block diagram. We can skip the off-line basis generation procedure as well as the online geometric feature extraction. This simplicity is a fruit of our visual feedback. The control objective is the asymptotic stabilization of the plant origin, that is, the steady horizontal surface, in the presence of occlusion. The initial surface condition at $t = t_0 > 0$ is prepared by applying the feedforward input (3) with $A_1 = A_2$ and $f_1 = f_2$ in the period $[0, t_0]$ to the steady horizontal surface at $t = 0$. Here, we set $(A_1, f_1) = (0.9, 0.285)$ for planar sloshing and $(A_1, f_1) = (1.50, 0.567)$ for nonplanar sloshing, and $t_0 = 15$ [s]. Just after the feedforward input ends, we start the LQG control minimizing the objective functions [8]

$$\Sigma_0^\infty \left(q_f \langle X(k), X(k)\rangle + r_f \langle U(k), U(k)\rangle\right)$$

for the LQ controller and

$$E[\mathrm{vec}(X(k) - \hat{X}(k))\mathrm{vec}(X(k) - \hat{X}(k))^{\mathrm{T}}]$$

for the Kalman filter with the estimated state $\hat{X}(k)$ against the zero-mean disturbances $V(k)$ and $W(k)$ such that

$$E[\mathrm{vec}(V(k))\mathrm{vec}(V(k))^{\mathrm{T}}] = q_e I_{m_x n_x}$$
$$E[\mathrm{vec}(W(k))\mathrm{vec}(W(k))^{\mathrm{T}}] = r_e I_N$$

in which $(q_f, r_f) = (0.008, 30.77)$ and $(q_e, r_e) = (0.001, 10)$ in case of $N = 4$ (4-output), and $(q_f, r_f) = (0.0142, 17.61)$ and $(q_e, r_e) = (0.001, 50)$ in case of $N = 64$ (64-output), respectively. These weights q_f, r_f, q_e, and r_e are searched so that the inputs at planar sloshing take the same value at $t = 15$ [s] between $N = 4$ and $N = 64$ for a fair comparison. First, in the absence of occlusion, the stabilization by the proposed control is discussed. Second, in the presence of occlusion which is a student's hand, the rejection and the attenuation of the whole occlusion effects are also discussed.

D. Control Experimental Results and Discussion

Fig. 13 shows the input component of $U(k)$ in case of $N = 4$, and Fig. 14 shows the corresponding output norm $\|Y(k)\|$ for nonplanar sloshing in the absence of occlusion. The dot (black) depicts the no control, and the cross (red) depicts the proposed control. The output norm $\|Y(k)\|$ grows until the initial time $t = 15$ [s] by the feedforward input and converges

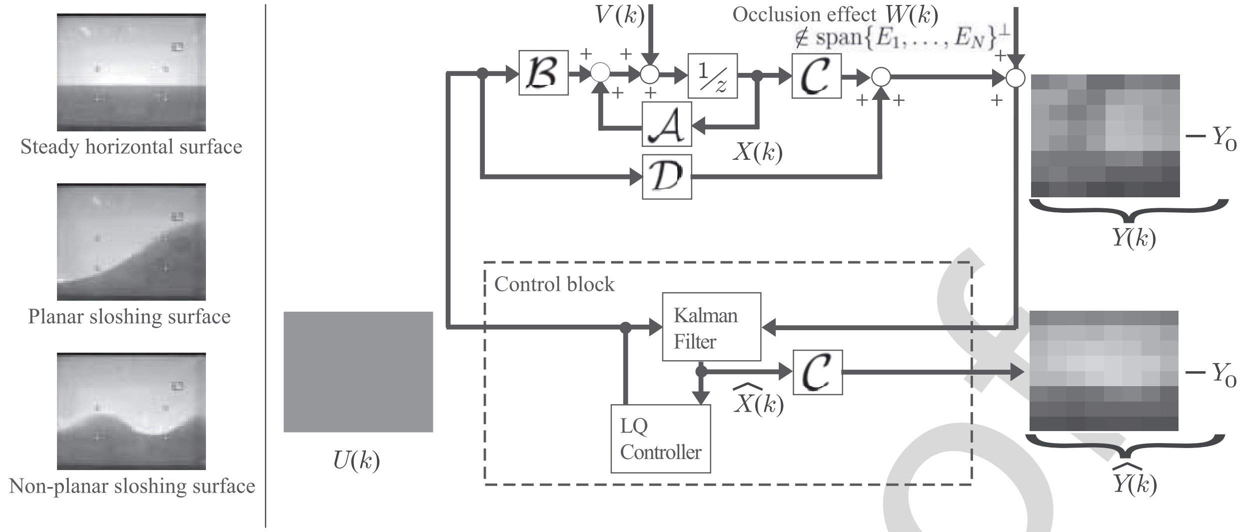

Fig. 12. Block diagram in which a part of the whole occlusion effects is rejected and the remaining part is attenuated.

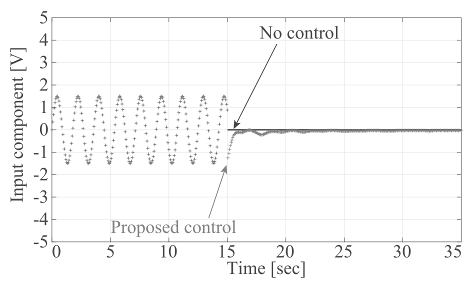

Fig. 13. Input component (without occlusion, $N = 4$).

TABLE I

MATRIX BASIS COMPARISON

Basis	Walsh $N=4$	Walsh $N=64$	POD $N=4$	POD $N=64$	No control
Cost (on-line)$^\diamond$	Low	High	Low	High	—
Cost (off-line)	**Low***	**Low***	High†	High†	—
Performance T_s	9.3	6.0	6.5	6.0	13.3

$^\diamond$The number of the basis elements (N)

*Two procedures (system identification, controller design)

†Three procedures (basis generation, system identification, controller design)

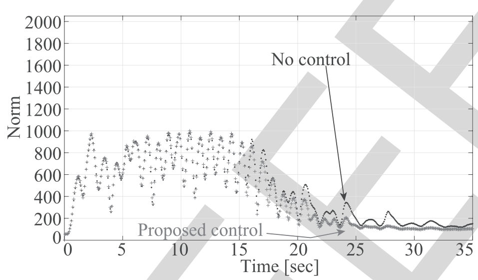

Fig. 14. Output norm (without occlusion, $N = 4$).

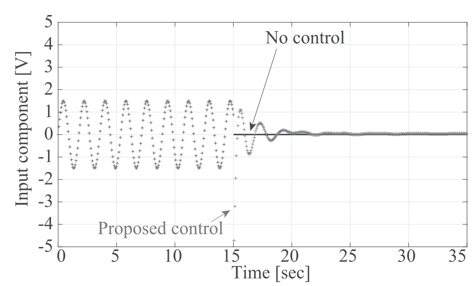

Fig. 15. Input component (without occlusion, $N = 64$).

to zero after $t = 15.0$ [s], as the input component converges to zero in the steady-state period. There is no input saturation. The convergence rate by the proposed control in case of $N = 4$ is slightly faster than that by the no control. The settling time by the proposed control in case of $N = 4$ is $T_s = 9.3$ [s] and that by the no control is $T_s = 13.3$ [s]. Here, the settling time T_s is a control performance introduced as the last time when the output norm is less than 20% of the maximum after we start the controls at $t = 15.0$ [s]. Note that Fig. 14 displays the high-resolution output for a fair comparison between $N = 4$ and $N = 64$.

Fig. 15 shows the input component of $U(k)$ in case of $N = 64$, and Fig. 16 shows the corresponding output norm $\|Y(k)\|$ for nonplanar sloshing in the absence of occlusion. There is no input saturation again. The settling time by the proposed control in case of $N = 64$ is $T_s = 6.0$ [s] and that by the no control is $T_s = 13.3$ [s] again. Especially, in the transient period $15 \leq t \leq 20$ [s], the convergence rate in case of $N = 64$ is much better than that in case of $N = 4$ successfully. In other words, the proposed control in case of $N = 4$ does not work well for nonplanar sloshing. This is because the space resolutions of the four basis elements $E_1^W, \ldots, E_4^W$ in Fig. 3 are lower than the others.

Table I summarizes the off-line and online experimental costs and the performance. The Walsh basis in case of $N = 64$ achieves the best performance. Here, $N = 64 (> 40)$ is very high so that the exchange of the Walsh basis for the alternative (POD) basis [9] can correspond to the change of basis and can

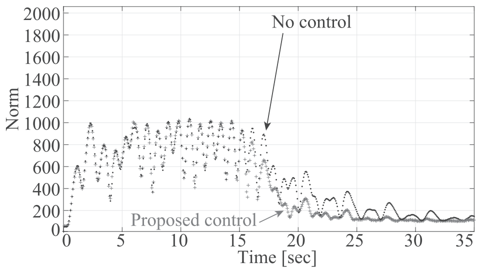

Fig. 16. Output norm (without occlusion, $N = 64$).

preserve the output norm

$$\|\mathcal{U} \circ Y(k)\| = \|Y(k)\|$$

for a unitary operator $\mathcal{U}$. This means that the alternative basis in case of $N = 64$ brings the same sigma plot in Fig. 10 and thus achieves the same performance $T_s = 6.0$ [s] experimentally in spite of the worse off-line experimental cost. On the other hand, the Walsh basis in case of $N = 4$ achieves the best experimental costs but has the worst performance than the alternative basis in case of $N = 4$. This is convincing, because the Walsh basis has no specific information about the plant block. In return, the Walsh basis can skip the off-line basis generation, and we do not have to handle the movies. The range of the basis design is increased successfully. The related paper [18] brings $T_s = 4.6-5.1$ [s] ($6.9-7.7$ [s] in the presence of occlusion) but is not fair here, because the design procedures of the image processing block are not systematic.

Fig. 17(a) shows the camera images for evaluation at the several discrete time by the proposed control in case of $N = 64$ in the absence of occlusion. The initial nonplanar surface tends to the plant origin. The validity is confirmed in the absence of occlusion.

On the other hand, Fig. 17(b) shows the camera images by the proposed control in case of $N = 64$ in the presence of occlusion. Fig. 17(c) shows the actual output $Y(k) + Y_0$, and Fig. 17(d) shows the estimated output $\hat{Y}(k) + Y_0$, which is calculated by the Kalman filter for the evaluation only and not for control. Here, the steady horizontal surface image Y_0 is added just for readability. Successfully, the initial nonplanar surface tends to the plant origin again. Also, by comparing Fig. 17(c) and Fig. 17(d), the occlusion effect in the output $Y(k)+Y_0$ is attenuated and almost disappeared in the estimated output $\hat{Y}(k)+Y_0$. Not only the plant origin but also the closed-loop origin is asymptotically stabilized.

Remark 3: Unlike conventional visual feedbacks, our visual feedback handles the occlusion in two steps without geometric features. In the first step, in the sense that the occlusion in Fig. 17(a) is projected to the occlusion effect that exists in $\mathbb{W} := \text{span}\{E_1^W, \ldots, E_{64}^W\}$, a part of the whole occlusion effects exists in $\mathbb{W}^\perp$ experimentally and is rejected. In the second step, the remaining part exists in $\mathbb{W}$ and is attenuated by the LQG control. The control performance in the presence of occlusion depends not only on the controller transfer function in the second step but also on the relation between the occlusion and the basis in the first step. There is no basis whose occlusion effect rejection performance is always better than the others for every possible occlusion.

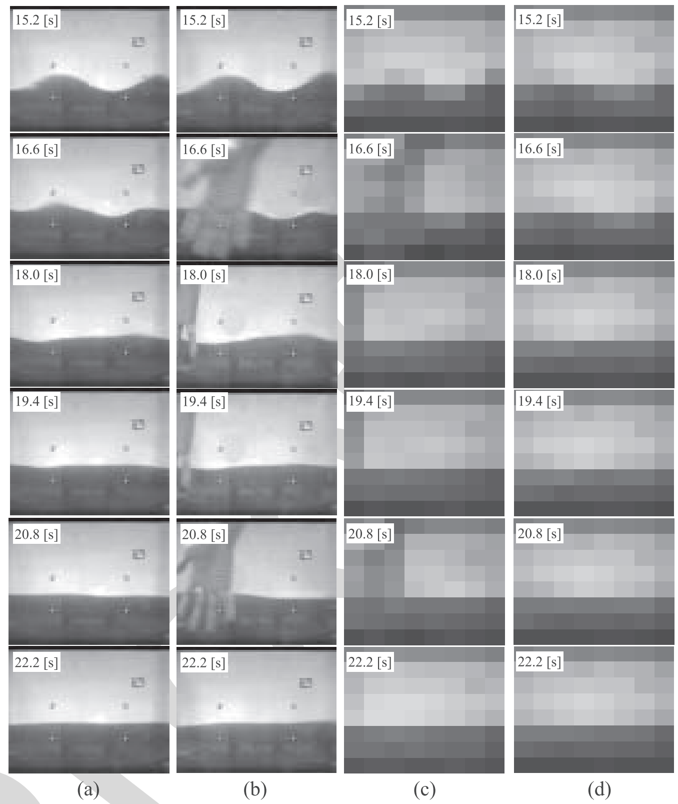

Fig. 17. Images and outputs. (a) Images (without occlusion). (b) Images (with occlusion). (c) $Y(k) + Y_0$ (with occlusion). (d) $\hat{Y}(k) + Y_0$ (with occlusion).

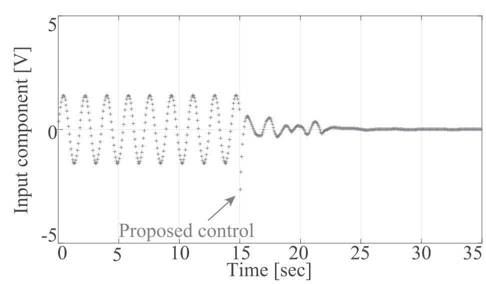

Fig. 18. Input component (with occlusion, $N = 64$).

Figs. 18 and 19 show the input component and the output norm in the presence of occlusion for nonplanar sloshing discussed in Fig. 17(a)–(c). These settling times are slightly larger than those without occlusion. Especially in the transient period $15 \leq t \leq 25$ [s], the existence of the occlusion is observed, but the input and output components tend to be zero in the steady-state period again. The validity is confirmed even in the presence of occlusion. As a demonstration, Fig. 20 shows the output norm $\|Y(k)\|$ against the input disturbance [the same chirplike input (3) for the system identification] instead of the output disturbance (the occlusion effect). Again, the proposed control is better than the no control.

Finally, let us discuss the robust stability analysis. This is also a demonstration that our visual feedback guarantees the closed-loop stability even in the presence of the (input multiplicative) uncertainty Δ and the occlusion effect W. Taking the extended block structure set: $\mathbf{\Delta} = \{\text{diag}(\Delta, \Delta_f) \mid \Delta \in \mathbb{C}, \Delta_f \in \mathbb{C}^{N\times N}\}$, it is known that the robust performance as

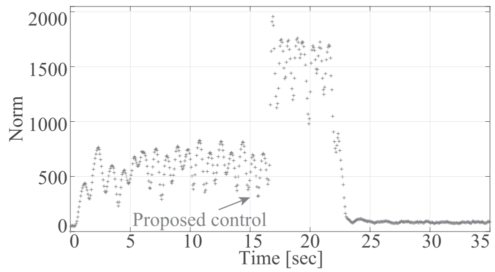

Fig. 19. Output norm (with occlusion, $N = 64$).

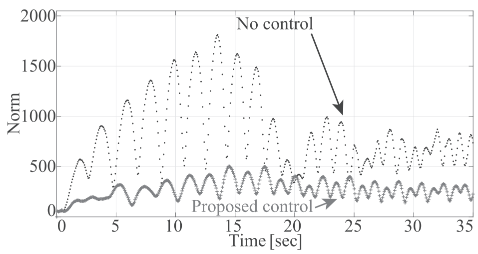

Fig. 20. Output norm (with input disturbance, $N = 64$).

well as the robust stability is evaluated by structured singular value (SSV): $\mu_{\mathbf{\Delta}}(G) = (\min\{\bar{\sigma}(\Delta) \mid \Delta \in \mathbf{\Delta}, \det(I - G\Delta) = 0\})^{-1}$ with a standard setting [12]

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \underbrace{\begin{bmatrix} -KS_oP & -KS_o \\ S_oP & S_o \end{bmatrix}}_{G} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$S_o = (I + L_o)^{-1}, \quad L_o = PK$$

where P is the nominal plant block from the input U to the output Y in the absence of W, K is the control block from the disturbed output Y to the input U of the perturbed plant block $P(1+\Delta)$, z_1 is the input U, z_2 is the disturbed output Y, $w_1 = \Delta z_1$, and w_2 is the output disturbance W. By the standard μ-analysis (MATLAB version 9.0) in the worst case of $N = 4$ and 64, in the presence of the uncertainty Δ with the observed maximum gain 2.087, one of the upper bounds of SSV is lower than 0.3612, which guarantees both of the robust stability for $|\Delta| \leq 2.087 < 1/0.3612$ and the robust performance $\|F_u(G, \Delta)\|_\infty < 0.3612$. The notation $F_u(\bullet, \bullet)$ denotes the upper linear fractional transformation. Unlike in the LQG controller design procedure, the zero-mean assumption is not needed, and the output of the uncertainty is the input disturbance. The tools (N4SID, LQG, and μ) in this brief are examples, and various other tools on the coordinate space are applicable to other dynamical systems (e.g., other continuum systems) on the matrix space as long as the input–output linearity exists in the sense of the system (1).

IV. Conclusion

For a visual feedback without geometric features, this brief suggests to apply a new special basis made by the Walsh functions to reduce the off-line experimental cost. The validity is confirmed experimentally against occlusion in nonplanar sloshing whose dynamics is not negligible. The range of the basis design is increased. The next work is a systematic basis generation to improve the input–output linearity as well as the occlusion effect rejection performance assuming that some *a priori* information about the occlusion is available.

Acknowledgment

The authors would like to thank the reviewers for the valuable comments, and they would also like to thank K. Nagai, S. Morozumi, and K. Lee for their experimental helps.

References

[1] G. Silveira and E. Malis, "Direct visual servoing: Vision-based estimation and control using only nonmetric information," *IEEE Trans. Robot.*, vol. 28, no. 4, pp. 974–980, Aug. 2012.

[2] C. Zang and K. Hashimoto, "GPU acceleration in a visual servo system," *J. Robot. Mech.*, vol. 24, no. 1, pp. 105–114, 2012. AQ:3

[3] J. Zhao, B. Song, N. Xi, L. Sun, H. Chen, and Y. Jia, "Non-vector space approach for nanoscale motion control," *Automatica*, vol. 50, no. 7, pp. 1835–1842, 2014.

[4] K. Deguchi and T. Noguchi, "Visual servoing using eigenspace method and dynamic calculation of interaction matrices," in *Proc. IEEE ICPR*, Aug. 1996, pp. 302–306.

[5] S. K. Nayar, S. A. Nene, and H. Murase, "Subspace methods for robot vision," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 750–758, Oct. 1996.

[6] A. Censi, S. Han, S. B. Fuller, and R. M. Murray, "A bio-plausible design for visual attitude stabilization," in *Proc. IEEE CDC*, Dec. 2009, pp. 3513–3520.

[7] S. Han, A. Censi, A. D. Straw, and R. M. Murray, "A bio-plausible design for visual pose stabilization," in *Proc. IEEE/RSJ IROS*, Oct. 2010, pp. 5679–5686.

[8] D. G. Luenberger, *Optimization by Vector Space Methods*. New York, NY, USA: Wiley, 1960.

[9] S. Sakai and M. Ando, "On the visual systems & control on matrix space," in *Proc. IEEE CDC*, Dec. 2014, pp. 2173–2178.

[10] K. Nagai and S. Sakai, "A visual feedback design on matrix space for a liquid sloshing experiment," in *Proc. SICE*, Sep. 2013, pp. 2088–2093.

[11] J. L. Walsh, "A closed set of normal orthogonal functions," *Amer. J. Math.*, vol. 45, no. 1, pp. 5–24, 1923.

[12] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

[13] K. G. Beauchamp, *Walsh Functions and Their Applications*. Norwell, MA, USA: Academic, 1975.

[14] W. K. Pratt, J. Kane, and H. C. Andrews, "Hadamard transform image coding," *Proc. IEEE*, vol. 57, no. 1, pp. 58–68, Jan. 1969.

[15] R. A. Ibrahim, *Liquid Sloshing Dynamics: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[16] T. Ikeda, M. Takashima, and Y. Harata, "Autoparametric resonances of elastic structures coupled with two sloshing modes in a square liquid tank," *ASME J. Comput. Nonlinear Dyn.*, vol. 8, no. 1, p. 011007, 2013.

[17] B. Robu, L. Baudouin, C. Prieur, and D. Arzelier, "Simultaneous H_∞ vibration control of fluid/plate system via reduced-order controller," *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 3, pp. 700–711, May 2012.

[18] S. Sakai and M. Sato, "Visual systems & control on polynomial space and its application to sloshing problems," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 6, pp. 2176–2187, Nov. 2014.

[19] K. Yano and K. Terashima, "Robust liquid container transfer control for complete sloshing suppression," *IEEE Trans. Control Syst. Technol.*, vol. 9, no. 3, pp. 483–493, May 2001.

[20] M. Grundelius and B. Bernhardsson, "Control of liquid slosh in an industrial packaging machine," in *Proc. Int. Conf. Control Appl.*, Aug. 1999, pp. 1654–1659.

[21] M. Peric, T. Zorn, O. El Moctar, T. E. Schellin, and Y.-S. Kim, "Simulation of sloshing in LNG-tanks," *ASME J. Offshore Mech. Arctirce Eng.*, vol. 131, no. 3, pp. 031101-1–031101-11, 2009.

[22] O. M. Faltinsen and A. N. Timokha, "On sloshing modes in a circular tank," *J. Fluid Mech.*, vol. 695, pp. 467–477, Mar. 2012.

[23] Q. Zang, J. Huang, and Z. Liang, "Slosh suppression for infinite modes in a moving liquid container," *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 1, pp. 217–225, Feb. 2015.

[24] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

PLEASE NOTE: We cannot accept new source files as corrections for your paper. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.

AQ:1 = Please provide the expansion for the acronyms “ARX and N4SID.”
AQ:2 = Please confirm whether the retention of the sentence “Note that our visual feedback is geometric feature less but not feature less” is correct.
AQ:3 = Please confirm the journal title, volume no., issue no., page range, and year for ref. [2].