Dynamics-based motion deblurring for a biologically-inspired camera positioning mechanism

Michael D. Kim, Student Member, IEEE and Jun Ueda, Member, IEEE

Abstract— This paper presents a new dynamics-based method for image processing in coordination with rapid ocular movement. A camera positioning mechanism with piezoelectric cellular actuators will be employed to demonstrate the effectiveness of this approach. There are a number of mechanisms for automatic camera positioning, but few have much in common with the human ocular positioning system. When a rapid point-to-point motion is created in a camera positioner, like human saccadic motion, the image sensor receives blurry images. Existing image techniques rely solely on obtained images, or use external sensors to estimate a blur kernel, or point spread function (PSF), for motion deblurring. Inspired by recent oculomotor studies, this paper proposes a method for estimating PSFs directly from the system dynamics of a camera positioning mechanism without motion sensors. The proposed method has been evaluated by using a single degree-of-freedom camera orientation system. Results are compared with conventional methods in terms of speed and accuracy.

I. INTRODUCTION

The human eye is a remarkable organ that detects light, allowing the brain to recognize the environment. Due to its limited field of view (FOV), human eyes need to change the orientation to scan the environment or to change the region of interest (ROI). Human eye movements are created by six surrounding extraocular muscles [1]. These six extraocular muscles consisting of four recti muscles and two oblique muscles create various eve motions in three degrees of freedom (DOF). Human eyes are known to create two representative movements: smooth pursuit and saccade. It is known that movement time is within 50ms for 10 degree saccade and the maximum saccade velocity is 250 degree per second [2]. Saccade is some of the fastest and most accurate movement made by human [3]. Saccade occurs much more rapidly than proprioceptive, vestibular, or visual feedback that can be returned to the brain [4]. This fact suggests that the saccade is completed in an open-loop manner in the control point of view [4].

A variety of camera positioning devices, inspired by the human ocular system, have been designed. Lan *et al.* have developed a two DOF camera-orientation mechanism by using rigid links and servo motors [5]. Lesmana *et al.* have developed a bio-inspired open-loop controller for fast eye movements by using DC servo motors [8]. Song *et al.* have developed an

M. D. Kim and J. Ueda are with Bio-Robotics Human Modeling Lab, George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0405, USA. (e-mail: mkim44@gatech.edu & jun.ueda@me.gatech.edu)

active binocular integrated system actuated by DC motors [9]. While these mechanisms successfully generated a rapid motion, heavy industrial servo motors have little in common with biological muscle actuation. In contrast, the human eye is positioned by means of antagonistic pairs of six extraocular muscles. There are several notable designs that orient the camera by means of antagonistic pairs of alternative actuators. Villgrattner et al. have developed a compact high-dynamic two DOF camera-orientation system driven by pairs of piezoelectric actuators [6]. Lenz et al. have designed a robot eve mechanism driven by pneumatic actuators [7]. Even though these designs have unique solutions to camera positioning, there is a discrepancy with muscle actuators. For example, the rigidity is an artifact of using traditional servo motors while the extraocular muscles are contractile and compliant. In addition, these actuators are controlled in closed-loop for a rapid motion while physiology study indicates that saccade is completed in an open-loop manner. Schultz et al. have developed a camera positioner driven by antagonistic pairs of compliant cellular actuators that are controlled by an open-loop, switching controller [10-11]. All aforementioned biologically inspired camera positioning robots are only focused on design and control of the mechanism but not on vision or image process. Although there was no vision processing for the previous studies, most of them employ a camera with a high frame rate over 100 frames per second (fps).

Contrary to a high-speed camera that was employed to the existing systems, the frame rate of the human eye is known to be around 24 fps. Therefore, during saccade or right after saccade is completed, the human eye will perceive blurred images. Physiology studies indicate that the brain predicts consequences of the eye movement so that motion smear can be reduced by a neural compensation mechanism using information of eye movements [13-14].

There are various approaches proposed in the area of computer vision for restoring degraded blurry images caused by motion. A point spread function (PSF) must first need to be estimated. The PSF is a blur kernel that describes the camera motion during exposure. Given the kernel, a blurry image can be reconstructed by using a standard deconvolution algorithm. Fergus *et al.* have attempted to estimate a PSF from a single blurred image and restored images using statics of an obtained image [15]. Multiple images have been used to remove blur in the images [16-17]. However, these approaches require a large amount of computation time for analyzing the blurry image.

This research was supported in part by NSF grant ECCS-0932208.



Figure 1. Camera orientation system

There are studies using different approaches by using addition sensors. Joshi *et al.* have used a gyro sensor and accelerometer to track a camera shake and to estimate a PSF for image deblurring [18]. Ben-Ezra et al. have used an extra high frame camera with a low resolution to track the camera motion and to estimate a PSF [19]. While these approaches were able to remove blur in images by using motion information, they also require a large amount of computational effort. Ben-Ezra *et al.*'s approach requires post image processes of multiple images obtained from a secondary high frame camera. Joshi *et al.*'s approach involves computations estimating the path of a camera shake from noisy gyro and accelerometer sensors.

This paper presents a biologically inspired estimation of the PSF by visual-motor coordination for image deblurring. The human brain uses information of eye movements to reduce motion smear [13-14]. This paper presents a method inspired by this observation. Understanding both mechanical properties and control architecture, PSFs can be estimated in an open-loop manner without the analysis of inherent image characteristics. This method directly estimates the PSF from the control system without any sensors resulting in less computational effort for image restoration. Although the ultimate goal of camera orientation system is a three degree-of-freedom system, the proposed method is tested and verified using single degree-of-freedom system as shown in Fig. 1.

II. DESIGN AND CONTROL OF THE CAMERA-ORIENTATION SYSTEM

A. Mechanical Design

The camera orientation system used in this study is shown in Fig. 1. A Logitech C270 HD webcam is used for image acquisition.

The moving platen is connected to a rod that transmits the force from antagonistic pairs of cellular actuators located on both sides. The cellular actuator consists of sixteen Lead Zirconate Titanate (PZT) stackactuators with deformable amplification mechanisms [12], [20]. The multi-layered strain



Figure 2. Single degree-of-freedom tilting motion of the camera orientation system

amplification structure amplifies the displacement of PZT from its extremely small strain. This mechanism exhibits zero backlash and noise-less operation since no gears or sliding mechanisms are used in the structures while still extremely fast movement is achieved. In addition, the nested compliant mechanism has been reported to resemble characteristics of human extraocular muscles [27].

The axis of the rod is assembled perpendicular but not orthogonal to the pivot axis. Therefore, when the rod is pushed or pulled by the antagonistic pairs of the cellular actuators, a moment is applied to the moving platen resulting in a tilting motion of the camera positioning system as shown in Fig. 2.

B. Mechanical Control

A total of 32 APA50XS (Cedrat corporation) PZT actuators are used to position the camera. The maximum displacement of the PZT stack actuator is 80 µm and it can be controlled continuously by adjusting the input voltage. The PZT actuator has hysteresis response requiring closed-loop control with additional sensing devices. In this study, however, the PZT is given only maximum input voltage of 150V or minimum input voltage of 0V in order to operate in an on-off manner so that the hysteresis response issue can be avoided. Since the system has a number of PZT stacks in an amplification mechanism to the maximum free displacement of 12.9mm, a discretized desired distance can be obtained by defining the numbers of on and off PZT actuators. In this approach, the cellular actuators share the principles in common with human muscles where a motion of the human muscles is a summation of motor units [10]. The camera positioning mechanism exhibits a single dominant frequency at 14.4 Hz with a damping ratio of 0.12. Therefore, the system can be represented as a linear second-order system given as:

$$G(s) = \frac{K}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{1}$$

where K is the residue, ω_n is the natural frequency, and ζ is the damping coefficient.

A saccade-like motion is an open-loop controlled point-to-point motion with 50ms settling time at the maximum angular velocity of 250deg/sec. In order for camera positioning mechanism to generate the saccade-like motion, quantized



Figure 3. (a) Time response of a rapid point-to-point motion of the camera orientation system ranging in various desired angles (b) Discrete switching commands to suppress vibration



Figure 4. Velocity profiles of a rapid point-to-point motion ranging in various desired angles given the discrete switching commands

input commands must be given at appropriate times to the PZT stacks with successful vibration suppression. This can be achieved by understanding the mechanical properties of the system such as the natural frequency and the damping ratio. By applying a phase-vector analysis of mechanism control [11], discrete switching commands can be given to the system as shown in Fig. 3(b). Figure 3(a) shows representative time responses of a point-to-point movement of the camera orientation system in various desired angles given integer values of amplitudes at appropriate times. The results show that the vibration has been suppressed and the settling time of each motion is within 60ms. If a step command is given to the system, however, the response shows considerable oscillation due to the lightly damped system. Figure 4 shows that the maximum peak velocity of the camera orientation system given the discrete switching commands is approximately 350 deg/sec. As a result, the speed of response of the camera positioning mechanism is comparable to that of the human ocular motion.

III. IMAGE ACQUISITION AND DEBLURRING

A. Exposure Windows

Images can be obtained by the camera in various situations when a motion is involved as shown in Fig. 5(d). Exposure Window I is that the camera first starts taking an image at a



Figure 5. (a) Exposure Window I. The image is obtained during partially stationary and in motion. The image is partially blurry. (b) Exposure Window II. The image is obtained in full motion (c) Exposure Window III. The image is obtained during partially stationary and in motion but fully blurred. (d) Exposure Windows of image acquisition by the bio-inspired camera orientation device

stable state, but soon the motion takes place during the image acquisition. This results in acquisition of a blurred image in partial region because data is not loaded to an image sensor simultaneously for all pixels. Exposure Window II is that an image is obtained in motion for full exposure window. Exposure Window III is similar to Exposure Window I where an image is obtained partially stationary and in motion, but the whole image is blurred. The period of image acquisition includes writing, reading and erasing of data. For Exposure Window III, the data is written during motion, but is read and erased when the motion is relatively steady-state. Exposure Window IV is a fully steady-state. The image acquisition in Exposure Window I is very different to the visual process in humans. For Exposure Window II, images are blocked in visual process of the human which is called saccadic masking [22]. Saccadic masking is the phenomenon in the brain where the brain blocks images with substantial motion blur in the course of the eye motion. Therefore, in this paper, images only in Exposure Window III are considered for deblurring.

B. Dynamics-based Kernel Estimation

A spatially-invariant blurred image \mathbf{B} can be represented by a convolution between a shift-invariant blur kernel \mathbf{K} and a latent sharp image \mathbf{I} plus \mathbf{N} :

$$\mathbf{B} = \mathbf{K} \otimes \mathbf{I} + \mathbf{N} \tag{2}$$

B, **I**, **N**
$$\in \mathfrak{R}^{m \times n}$$
, **K** $\in \mathfrak{R}^{l \times l}$

where \otimes is a convolution operator.

A blur kernel, or a PSF, represents a trajectory of the motion with energy levels. In order to perform deblurring, it is important to estimate a blur kernel or PSF accurately. To perform deblurring, it is important to estimate a blur kernel or PSF accurately. In the literature, PSFs are obtained by the analysis of a single image, requiring a large calculation time [15], [17]. Another approach is to use an additional high-speed camera. However, the hardware tends to be bulky [18-19]. This paper proposes the estimation of PSFs from the system dynamics of the camera orientation system. Excluding the case where the system has large external perturbations, the motion of a robotic camera positioning system is predictable since the dynamics of the systems and actuation commands are known and given. This predictability of the system allows for sensorless estimation of PSFs. This approach is inspired by the observation that the brain seems to use eye movements to reduce motion smear [13-14].

Since the camera positioning mechanism can be modeled as a linear, time-invariant second order system, the time response of the single degree-of-freedom camera orientation system can be given as:

$$\theta(t) = \sum_{i=1}^{n} A_i \cdot \frac{\omega_n}{\sqrt{1-\zeta^2}} e^{-\zeta \omega_n(t-t_i)} \sin(\omega_n \sqrt{1-\zeta^2} (t-t_i))$$
⁽³⁾

where A_i is the amplitude of the i^{th} impulsive input, t_i is the time of given i^{th} impulsive input, n is the number of amplitudes given to the system, ω_n is the natural frequency, and ζ is the damping coefficient.

A PSF can be given as

$$k(x, y) = h_{(x, y)} \tag{4}$$

where *h* is an energy level at the location (x, y) of a pixel.

The pixel of the kernel (x, y) is

$$(x, y) = L \int_{t_a}^{t_b} \theta(t) dt \hat{x}$$
 (5)

where L is a conversion factor dependent on the size of the blur kernel.

and the energy function is

$$h_{(x,y)} = \frac{t_b - t_a}{t_{aca}} \tag{6}$$

where t_{aca} is the total time of single image acquisition.

Thus, the energy is proportional to the time remaining at the location (x, y). The PSF is an energy distribution function for which the energy conservation constraint must hold:

$$\iint k(x, y) dx dy = 1 \tag{7}$$

Given the mechanical properties and control commands of the system, therefore, a discretized PSF can be estimated without the use of sensors. This results in fast estimation of the PSF because it is estimated simultaneously with the image acquisition. On the contrary, image-based methods require a large amount of computation time for estimation and sensor-based methods also require a large amount of computational effort to convert data to the path of a camera shake.

C. Deconvolution

After estimating the PSF, image deblurring can be achieved by means of a deconvolution process. In this paper, a widely-used Richardson-Lucy deconvolution method is chosen. This method is known to be robust against noise. It produces deblurring images with high quality in presence of high noise levels. It recovers the latent image iteratively with an estimated PSF [23]. For the deconvolution process, MATLAB's deconvlucy.m function can be used

$$\mathbf{I}^{n+1} = \left(\frac{\mathbf{B}}{\mathbf{I}^n \otimes \mathbf{K}} \otimes \hat{\mathbf{K}}\right) \mathbf{I}^n \tag{8}$$

where **B** is the obtained image, **K** is the blur kernel, \mathbf{K} is the flipped blur kernel, and \mathbf{I}^n is the recovered image at *n* th iteration.



Figure 4. Visual-motor coordination for a biologically inspired estimation of the PSF

IV. EXPERIMENTAL SETUP

The PSF can be estimated in an embedded controller, NI cRIO-9024, in real-time as shown in Fig. 6. First an image is obtained from the Logitech C270 and desired angle is defined from the inverse kinematics. Second, open-loop commands are generated in the embedded controller and given to the camera orientation system. Lastly, the system produces a rapid point-to-point motion shown in Figs. 3 and 4. The PSF for image deblurring is estimated in the embedded controller in an open-loop manner. An image obtained in phase III, discussed in the previous section, is used for deconvolution. The images are obtained 1280x960 RGB 24 format in DirectShow bus at maximum 30 fps via USB 2.0 communication.

V. RESULTS AND DISCUSSION

Figure 7 and 8 show the results of the proposed image deblurring method. The results are compared with other

well-known algorithms: methods proposed by Xu et al. [24], Goldstein et al. [25], Shan et al. [26], and Fergus et al. [15]. The top row shows full images and bottom row shows cropped local images to clearly show the results. The blurred images are obtained in Exposure Window III. The images in Fig.9 are taken at 25 fps similar to a typical *frame speed* of the human eves [28]. The images in Fig. 10 are taken at 15 fps due to low ambient light. The results show an improvement over the input blurry images. It can be seen that the results are better than Xu et al.'s method, Shan et al.'s method and Fergus et al.'s method. The Shan et al. and Fergus et al. results show large ringing artifacts. The Xu et al. sharpens the original image but there are still residual ringing artifacts. The results of Goldstein et al. method are comparable. However, the output images from all methods are still locally blurry. The reason is that the image blur is introduced due to a rapid motion of the camera with short exposure window while the previous studies have focused on the image blur caused by long exposures due to low light. Also, a spatial-invariant blur kernel is used for image restoration.

TABLE I. CONDITIONS FOR COMPARISON OF COMPUTATOIN TIME

	Kernel Size	Image Size	Software
Xu et al. [24]	19x19	800x600	Complied Executable*
Goldstein et al. [25]	21x21	800x600	MATLAB
Shan et al. [26]	27x27	800x600	Compiled Executable*
Fergus et al. [15]	21x21	800x600	MATLAB
Proposed Method	21x21	800x600	MATLAB

(*: Complied executable files are distributed by the Authors)

The output images shown in Fig. 8 are deblurred, but slightly noisy due to an existence of dark current noise in the captured images. It can be seen that the methods of Xu et al., Shan et al., and Fergus et al. are not robust to dark current noise. Figure 9 shows a comparison of computation times. The



Figure 5. Deblurring results and comparisons. Images are taken at 25fps. Top row: full image. Bottom row: cropped local image. The cropped region is specified in the full blurry image (a) The original blurred input image (b) Xu *et al.* [24] (c) Goldstein *et al.* [25] (d) Shan *et al.* [26] (e) Fergus *et al.* [15] (f) Proposed method



Figure 6. Deblurring results and comparisons. Images are taken at 15fps. Top row: full image. Bottom row: cropped local image. The cropped region is specified in the full blurry image. Images are taken under low ambient light. (a) The original blurred input image (b) Xu *et al.* [24] (c) Goldstein *et al.* [25] (d) Shan *et al.* [26] (e) Fergus *et al.* [15] (f) Proposed method. The output images of some methods are degraded



Figure 7. Comparison of computation times

computation times include estimation of a PSF and deconvolution. A total of 10 trials are executed for each algorithm. The results show that the proposed method is the fastest since the PSF is estimated from the embedded motion controller in a parallel process while those of others are estimated by the analysis of images. The computation time of the proposed method could be decreased by using a non-iterative deconvolution method.

VI. CONCLUSION AND FUTURE WORK

This paper has presented a biologically inspired estimation of the spatial-invariant PSF. The PSF is estimated in an open-loop manner based on the dynamics of the system with no motion sensors. Blurry images were obtained in Exposure Window III. The blurry image was restored by the deconvolution process in knowledge of the estimated PSF. The results show that the proposed method is effective for image deblurring caused by fast motion. In addition, the computation time is faster than other approaches. Future work will involve an estimation of the spatial-variant PSF from mechanical properties and control architecture, fast generation of panoramic images, and visual servoing by using a multiple DOF camera positioner. Once the PSF is estimated, the controller should be able to determine an ideal profile to position the camera and exposure times to take images in order to re-use the estimated PSF.

ACKNOWLEDGMENT

The authors wish to express their thanks to Dr. Joshua Schultz for the valuable work on the camera positioning hardware and control.

REFERENCES

- [1] R. H. S. Carpenter, *Movements of the eyes*, 2nd ed. London: Pion, 1988.
- [2] C. M. Harris and D. M. Wolpert, "Signal-dependent noise determines motor planning," *Nature*, vol. 394, pp. 780-784, 1998.
- [3] R. J. Leigh and C. Kenard. "Using saccades as a research tool in the clinical neurosciences," *Brain*, vol. 127, no. 3, pp. 460-477, 2004.
- [4] D. A. Robinson, "The mechanics of human saccadic eye movement," J. of Physiology, vol. 174, pp. 245-264, 1964.
- [5] Y.-C. Lee, C.-C. Lan, C.-Y. Chu, C.-M. Lai and Y.-J. Chen, "A Pan-Tilt Orienting Mechanism with Parallel Axes of Flexural Actuation," *IEEE/ASME Trans. Mech.*, vol. 18, no.3, pp.1100-1112, June 2013.

- [6] T. Villgrattner and H. Ulbrich, "Design and control of a compact high–dynamic camera–orientation system," *IEEE/ASME Trans. Mech.*, vol. 16, no. 2, pp. 221-231, April 2011.
- [7] A. Lenz, S. R. Anderson, A. G. Pipe, C. Melhuish, P. Dean and J. Porrill, "Cerebellar-Inspired Adaptive Control of a Robot Eye Actuated by Pneumatic Artificial Muscles," *IEEE Trans. Sys., Man and Cybernetics: Part B*, vol. 39, no. 6, pp. 1420-1433, Dec. 2009.
- [8] M. Lesmana and D. K. Pai, "A Biologically Inspired Controller for Fast Eye Movements," *in Proc. IEEE Int. Conf. Robotics and Automation*, 2011, vol., no., pp.3670-3675.
- [9] Y. Song and X. Zhang, "An Active Binocular Integrated System for Intelligent Robot Vision," in Proc. IEEE Int. Conf. Intelligence and Security Info., 2012, vol., no., pp. 48-53.
- [10] J. Schultz and J. A. Ueda, "A camera positioner driven by muscle-like actuation," in Proc. IEEE Int. Conf. Biomed. Robot. Biomechatron., 2012, pp. 719–724
- [11] J. Schultz and J. Ueda, "Experimental verification of discrete switching vibration suppression," *IEEE/ASME Trans. Mech.*, vol. 17, no. 2, p. 298-308, 2012.
- [12] A. Ehud and A. Amos, "Seeing via miniature eye movements: a dynamic hypothesis for vision," *Frontiers in Computational Neuroscience*, vol. 6, no. 89, pp. 1-27, 2012.
- [13] J. Tong, S.B. Stevenson, and H.E. Bedell, "Signals of eye-muscle proprioception modulate perceived motion smear," *Journal of Vision*, vol. 8, no. 14, pp. 7.1-6, Oct. 2008.
- [14] M. Rolfs, D. Jonikaitis, H. Deubel and P. Cavanagh, "Predictive remapping of attention across eye movements," *Nature Neuroscience*, vol. 14, pp.252-256, 2010.
- [15] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, Removing camera shake from a single photograph," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 787–794, 2006.
- [16] M. Shi and J. Y. Zheng, "A slit scanning depth of route panorama from stationary blur," in Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition, 2005, vol. 1, pp. 1047–1054.
- [17] L. Bar, B. Berkels, M. Rumpf and G. Sapiro, "A variational framework for simultaneous motion estimation and restoration of motion-blurred video," *in Proc. IEEE 11th Int. Conf. Computer Vision*, 2007, pp. 1–8.
- [18] N. Joshi, S. B. Kang, C. L. Zitnick, and R. Szeliski, "Image deblurring using inertial measurement sensors," *ACM Trans. Graph.*, vol. 29, no. 4, pp. 1–9, 2010.
- [19] M. Ben-Ezra and S. Nayar, "Motion-based Motion Deblurring," *IEEE Trans. Pattern Analy. and Mach. Intell.*, vol. 26, no. 6, pp. 689-698, 2004.
- [20] J. Ueda, T. Secord, and H. Asada., "Large effective-strain piezoelectric actuators using nested cellular architecture with exponential strain amplification mechanisms," *IEEE/ASME Trans. Mech.*, vol. 15, no. 5, pp. 770–782, Oct. 2010.
- [21] J. W. Miskin and D. J. C. MacKay, "Ensemble learning for blind image separation and deconvolution," in *Advances in Independent Component Analysis*, M. Girolami, Ed. New York: Springer-Verlag, July 2000.
- [22] Matin, Ethel. "Saccadic suppression: a review and an analysis." Psychological bulletin 81.12 (1974): 899.
- [23] D. Fish, A. Brinicombe, E. Pike, and J. Walker, "Blind Deconvolution by Means of the Richardson-Lucy algorithm," J. Optical Soc. Am., vol. 12, 1995.
- [24] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proc. ECCV*, 2010, pp. 157–170
- [25] A. Goldstein and R. Fattal, "Blur-Kernel estimation from spectral irregularities", in *Proc. ECCV*, 2012, pp.622-635
- [26] Q. Shan, J. Jia, and A. Agarwala, "High-Quality Motion Deblurring from a Single Image," ACM Trans. Graphics, vol. 27, no. 3, 2008.
- [27] Jun Ueda, "Piezoelectrically Actuated Robotic End-effector with Strain Amplification Mechanisms," Advanced Mechatronics and MEMS Device, Microsystems, vol. 23, pp. 25-52, Edited by Dan Zhang, Springer, ISBN 978-1-4419-9984-9, 2012.
- [28] F. Kuester and T. C. Hutchinson, "A virtualized laboratory for earthquake engineering education," *Comput. Appl. Eng. Edu*, vol. 15, pp. 15-29, 2007