# Dynamic Photometric Stereo Method using Multi-tap CMOS Image Sensor

Takuya Yoda[1], Hajime Nagahara[1], Rin-ichiro Taniguchi[1],
Keiichiro Kagawa[2], Keita Yasutomi[2], Shoji Kawahito[2]
1 Graduate School of Information Science and Electrical Engineering, Kyushu University
744, Motooka, Nishi-ku, Fukuoka, 819-0395 Japan
2 Research Institute of Electronics, Shizuoka University
3-5-1 Johoku, Nakaku, Hamamatsu, Shizuoka 432-8011, Japan
E-mail:yoda@limu.ait.kyushu-u.ac.jp

**Abstract** Photometric stereo enables the estimation of surface normals from images that were captured using different known lighting directions. The classical photometric stereo method requires at least three images to determine the normals of a given scene. This method therefore cannot be applied to a dynamic scene, because it is assumed that the scene should remain static while the required images are captured. We present a dynamic photometric stereo method to estimate the surface normals in a dynamic scene. We use a multi-tap CMOS image sensor to capture the input images for the photometric stereo method. The image sensor can divide the electrons from the photodiode of a single pixel into different taps of exposures, and can therefore capture multiple images under different lighting conditions with almost the same timing. We implemented a prototype camera that was synchronized with a lighting system, and subsequently realized photometric stereo of a dynamic scene.

**Keywords: Computational Photography, Photometric Stereo, surface normal, CMOS image sensor**

## 1. Introduction

Acquisition of 3D information has drawn considerable research attention in recent years, and various methods have been proposed using devices such as image sensors and laser scanners. Woodham [1] first proposed the photometric stereo method, which can estimate the dense surface normals from a set of images under different lighting directions. This method uses three images and known lighting conditions. If the object is assumed to be the Lambertian model, we can obtain the intensity images from the following equation:

$$\begin{bmatrix} I_n^1 \\ I_n^2 \\ I_n^3 \end{bmatrix} = N \begin{bmatrix} L^1 \\ L^2 \\ L^3 \end{bmatrix}, \qquad (1)$$

where $L^i$ represents the directional vectors of light sources. We obtained the camera and light source positions prior to calibration; $L^i$ was taken from the relative positions and $N$ represents the surface normals matrix. Therefore, we can obtain the surface normals from the following equation (2). $L^\dagger$ represents a pseudo-inverse matrix of the light source positions.

$$N = \begin{bmatrix} I_n^1 \\ I_n^2 \\ I_n^3 \end{bmatrix} \begin{bmatrix} L^1 \\ L^2 \\ L^3 \end{bmatrix}^\dagger . \qquad (2)$$

However, the photometric stereo method has two strong limitations. First, the method requires at least three input images acquired under different lighting conditions. Second, both the camera and the scene should remain static while all three input images are captured. These limitations stem from the assumptions made in the algorithm that the pixels of the three images should correspond and that the intensities of these corresponding pixels are only affected by changes in the lighting. Under these assumptions, the photometric stereo method linearly estimates a normal vector for each pixel from the intensities of the three images. Classical photometric stereo therefore cannot be applied to a dynamic scene, because the pixels at the same positions in the images are not the same pixels that correspond to the object position if the scene is dynamic.

In this paper, we propose a dynamic photometric stereo method using a multi-tap CMOS image sensor. The image sensor can divide the electrons from a photodiode into multiple exposures in a single pixel and obtain multiple images with almost the same timing. We built a prototype camera and lighting system and synchronized them to capture input images under different directional lighting. We can therefore estimate the normal map of a dynamic scene using photometric stereo from the captured images.

## 2. Dynamic photometric stereo method using Multi-tap CMOS image sensor

CMOS image sensors are increasingly popular in commercial products, because they feature system-on-chip integration and low power consumption. Regular CMOS image sensors for photography use a single photodiode in each pixel. The photodiode converts photons into electrons via the photoelectric effect. These electrons are then charged to a storage diode during the exposure time. The charged electrons are subsequently read out and form a single digital image such that the intensity of each pixel corresponds to the electrons and the number of photons.

Fig.1 shows a comparison of the timing diagram of the exposure and readout times. In general, high-speed cameras can capture several images in the time taken by a standard camera to capture one image. However, the high-speed camera exposure for each image must be short, and images thus have lower SNRs.

A multi-tap CMOS sensor has been proposed in the literature [2]. This sensor has multiple floating diffusions(FDs) that can split the electrons generated on the photodiode of a single pixel.

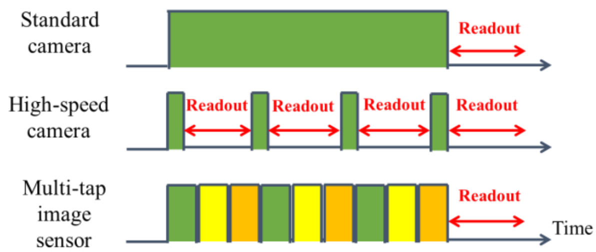


Fig.1 Timing diagram comparison of standard camera, high-speed camera, and multi-tap CMOS image sensor.

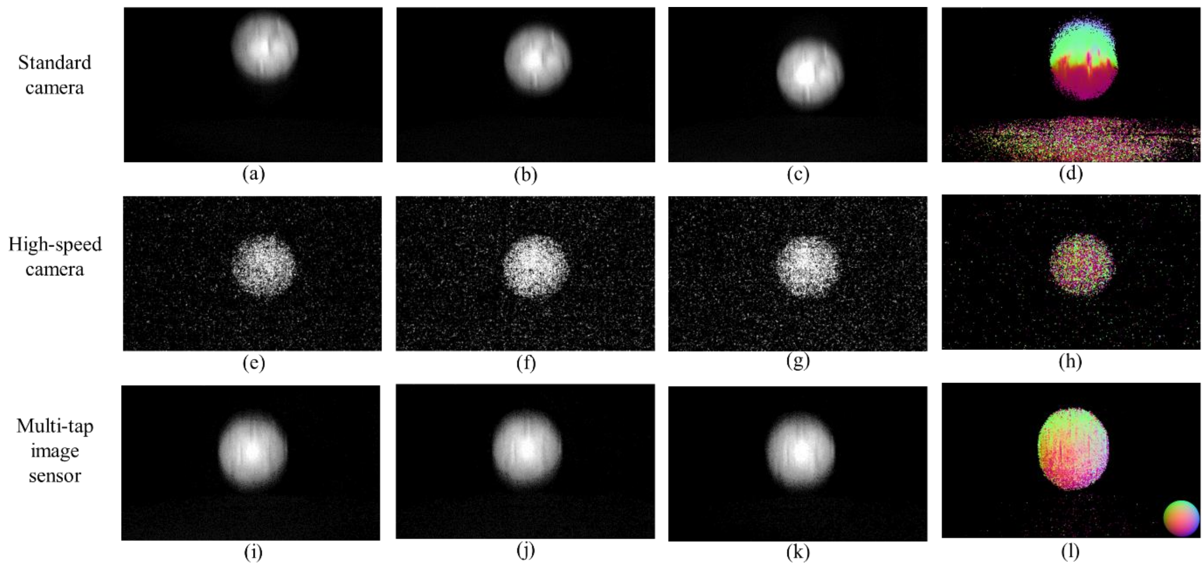


Fig.3 Photometric stereo results comparison of a standard camera, a high-speed camera, and the multi-tap image sensor. We captured the images of a falling ball.

We use the multi-tap CMOS image sensor [2], and it has three FDs. The sensor can charge the electrons to each FDs one by one and iterate this process multiple times, we obtain multiple partitions of the exposures as Fig.1. We can obtain three different images as integration of the three different colored regions of the exposures in the readout process at the end of exposures.

We propose a camera and lighting system that enables input image capture for photometric stereo in a dynamic scene. We use the multi-tap sensor [2] in the proposed system. There are three controlled lights and these lights are completely synchronized to the sensor exposures. Fig.2 shows the synchronization between light sources and the sensor exposures. Each light beam is emitted as multiple iterative pulses. The duration and timing of each light beam is completely synchronized to the corresponding tap of the exposure. The light emission durations and exposure times of the FDs can be reduced to microseconds. This is fast enough to ignore the scene dynamics, but is still sufficient to light the scene separately using several different lights. We obtain the three images with the different lighting angles required for the photometric stereo algorithm at almost the same time.

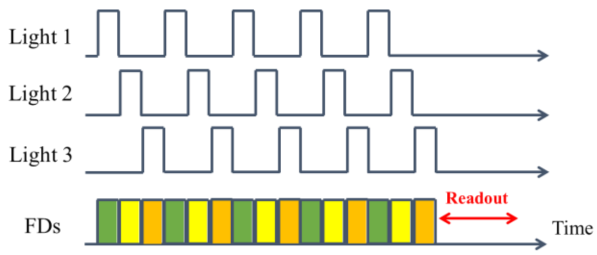


Fig.2 Synchronization timing chart between the light sources and exposure times of different FDs.

## 3. Experimental results

We compared the proposed method with the previous photometric stereo method. We captured images of a falling ball to act as the target dynamic scene. Fig.3(a)-(c) show the input images captured using the standard camera, while Fig.3(e)-(g) show those captured by the high-speed camera and Fig.3(i)-(k) show those of the multi-tap camera. We used the same camera to make the comparison, but we set the different exposure settings, 12 ms exposure for one image of the standard, 0.2 ms exposure for one image of the high-speed, and 4 ms exposure (20 iterations of 0.2 ms) for one image of the multi-tap. Position differences can be seen in the images of Fig.3(a)-(c) because the frame rate is insufficient to ignore the speed of the falling ball. As a result, Fig.3(d) shows that the normal map has a wrong estimation that comes from poor correspondence. Conversely, Fig.3(e)-(g) and Fig.3(i)-(k) did not show these differences in their images, although the ball is actually falling. Fig.3(h) and 3(l) show the correct normal maps, but Fig.3(h) shows the estimation errors because the input images of Fig.3(e)-(g) contain a lot of noise that comes from the short exposure times of the high-speed camera. We can therefore confirm that our proposed method using the multi-tap CMOS image sensor can apply the photometric stereo method to dynamic scenes and has advantages when compared with the previous high-speed camera approach.

## 4. Conclusion

In this paper, we have presented a dynamic photometric stereo method using a multi-tap CMOS image sensor. As the experimental results shows, this method can estimate the surface normals of a dynamic scene.

## References

[1] R. J. Woodham, "Photometric method for determining surface orientation from multiple images", Optical Engineering, vol.19, no.1, pp.139-144, 1980.

[2] S.-M. Han, T. Takasawa, K. Yasutomi, S. Aoyama, K. Kagawa, and S. Kawahito, "A time-of-flight range image sensor with background canceling lock-in pixels based on lateral electric field change modulation", IEEE Journal of the Electron Devices Society, vol. 3, pp. 267-275, 2015.