Computational Photography using Programmable Sensors

Hajime Nagahara[1]
1 Institute for Datability Science, Osaka University
2-8, Yamadaoka, Suita, Osaka, 565-0871, Japan
E-mail:nagahara@ids.osaka-u.ac.jp

Abstract A regular camera has an image sensor which has electric shutter called as global shutter. Global shutter expose all of the pixels at the same timing and captures moment of a scene. Recently, new types of programmable sensors which uniquely controls the exposure pattern of the pixels. This talk present examples of computational photography researches utilizing the programmable sensors.

Keywords: Programmable sensor, Compressive video sensing, Photometric stereo

1. Overview

A regular camera has an image sensor which has global shutter which exposure all of the pixels at the same time. We can get the sill image at the moment of the scene, and also obtain a video as a sequence of the images. Recently, there are some special sensors [1,2] which can more uniquely control the exposure patterns in space and time. These programmable sensor has been designed for the specific purpose and applications such as high dynamic imaging or time-of-flight for 3D measurement. However, these unique function or programmability of the exposure timing has a potential for the applications. This talk present the examples of the applications, such as compressive video sensing, privacy protected surveillance, and photometric stereo by using programmable sensors.

The sensor [1] has a row and column wise reset and transfer line. It can control these signals for controlling pixels wise exposures in each 8x8 block. We can utilize this programmable exposure function for compressive video sensing application[3,4] for generating x16 higher frame-rate video from the each coded exposed captured frame. We have discussed how we effectively design the coded exposure pattern for compressive video by using machine learning framework in the paper[4]. The sensor[1] is not fully controllable and could not change the pattern on runtime. If we have an ideal programmable sensor, we can change the region of interest or region of uninterest to be exposed in the scene adaptively to the context. We showed the concept, that camera do not optically capture the facial region but can capture the other, to solve a privacy problem of surveillance camera[5].

Photometric stereo is well known method to recover the detailed of surface of 3D object[6]. However, it is also well know that photometric stereo cannot be applied to the dynamic scene, since it requires at least three images with lit by different directional lights. This was the fundamental problem of the photometric stereo. A multi-tap sensor which is commonly used for time-of-flight camera has multiple exposures in the single pixels[2]. The sensor can take multiple image at almost the same timing. We used this sensor for the images under the different directional lighting which photometric stereo requires and realize to reconstruct the surface normal for the dynamic scene. This research solve the previus limitation of the photometric stereo and also show the new application of multi-tap sensor.

We have been studing about a lot of attempts of the researches to utilize the flexibllity and fuction of programmable sensor and looking for the new killer applications of these sensors. If you need the details, please see the seriease of researches in the reference.

References

[1] Hamamatsu Photonics K.K., “Imaging device”, Japan patent, JP2015-216594A, 2015-12-03.

[2] S. M. Han, T. Takasawa, K. Yasutom, 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 Charge Modulation, IEEE Journal of the Electron Devices Society, Vol. 3, No. 3, pp. 267-275, (2015).

[3] T. Sonoda, H. Nagahara, K. Endo, Y. Sugiyama, and R. Taniguchi, "High-Speed Imaging using CMOS Image Sensor with Quasi Pixel-Wise Exposure", ICCP (2016).

[4] M.Yoshida, A. Torii, M. Okutomi, K. Endo, Y. Sugiyama, R. Taniguchi, H. Nagahara, “Joint optimization for compressive video sensing and reconstruction under hardware constraints”, European Conference on Computer Vision(ECCV2018), Munich, Germany, Sep., (2018).

[5] Y. Zhang, Y. Lu, H. Nagahara, R. Taniguchi, “Anonymous camera for privacy protection”, International Conference on Pattern Recognition(ICPR2014), pp. 4170-4175, Stockholm, Sweden , Aug., (2014).

[6] R. J. Woodham: Photometric method for determining surface orientation from multiple images, Optical Engineering, Vol. 19, No. 1, pp. 139-144, (1980).

[7] T. Yoda, H. Nagahara, R. Taniguchi, K. Kagawa, K. Yasutomi, and S. Kawahito, “Dynamic Photometric Stereo Method using Multi-tap CMOS Image Sensor”, IEEE International Conference on Pattern Recognition, (2016).

[8] T. Yoda, H. Nagahara, R. Taniguchi, K. Kagawa, K. Yasutomi and S. Kawahito, “The Dynamic Photometric Stereo Method Using a Multi-Tap CMOS Image Sensor”, MDPI Sensors, Vol. 18, No. 3, Mar., (2018).

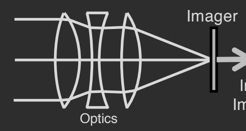
Computational photography using a programmable sensor

Hajime Nagahara
Institute for Datability Science
Osaka University

Imaging pipeline of digital camera

Scene → Optics → Imager → Image → Understanding

Hardware | Software

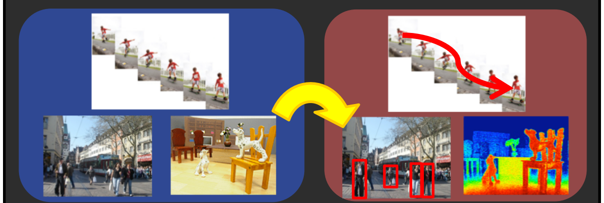


2

Imaging pipeline of digital camera

Scene → Optics → Imager → Image → Understanding

Hardware | Software



Captured image
(High resolution, high frame rate)

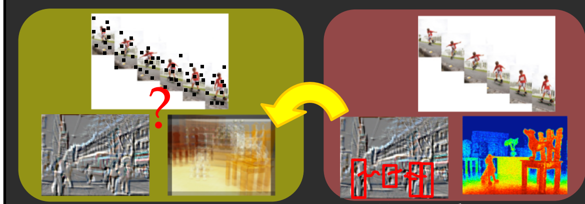
Output
(Motion, Object recognition, depth)

It is difficult to realize them by an algorithm improvement.

Computational Photography

Scene → Optics → Imager → Image → Understanding

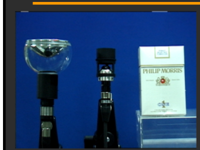
Processing based imaging
Task oriented image sensing



What is the optimal sensing?
(Coding, multiplexing, sparse sensing)

Target applications
(High speed video, depth)

Computational Photography

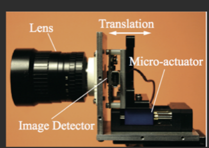


Omnidirectional camera

Hybrid camera

Coded aperture camera

Pixel coding camera



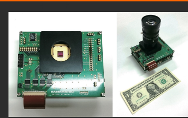
Focus sweep camera

Anonymous camera

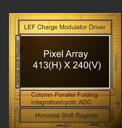
CP + Imager

Scene → Optics → Imager → Image → Understanding

- Programmable exposure CMOS
  - Controllable to exposure timing at each pixel
  - Collaboration with Hamamatsu photonics
  - Compressive video sensing, privacy protection for surveillance
- Multi-tap CMOS
  - Capable to multi-backet exposures
  - Collaboration with Shizuoka Univ.
  - Time of Flight, Photometric stereo



[Prototype camera]



[Kagawa et al. ISSCC14]

Space time sampling in Cameras

Still camera
- 8688 x 5792 pixels
- Less than 1 frames/sec.

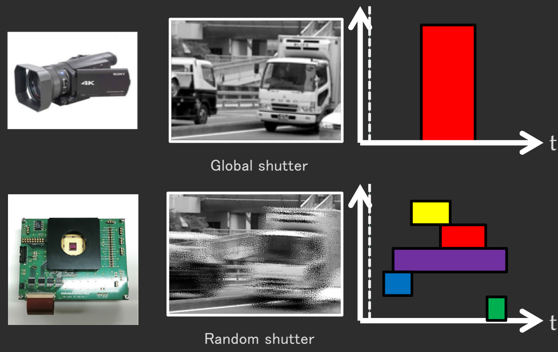
Video
- 3840×2160 pixels (4KHD)
- 60 frames / sec.

Data transfer rate = Resolution × frame rate

High resolution ⟷ High frame rate

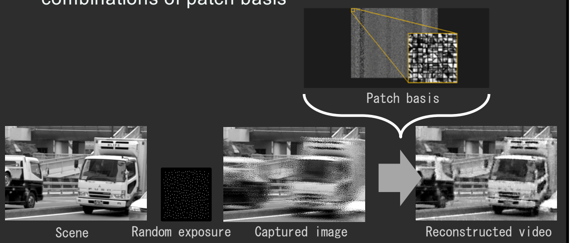
Trade-off

Random shutter

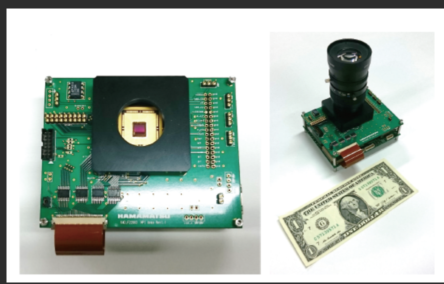


Sparse reconstruction for compressive video

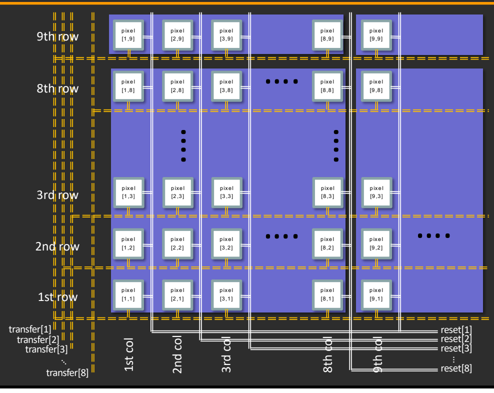
- Recovering a video from the captured image with random exposures
- Assuming sparsity of a scene and recovering it as combinations of patch basis



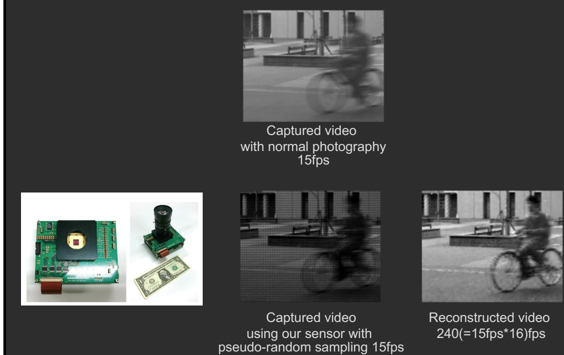
Pixel coding CMOS



Pixel coding CMOS



Quasi-random exposure for CS

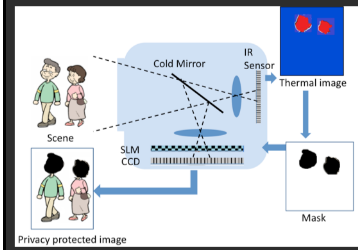


[Sonoda+ ICCP2016]

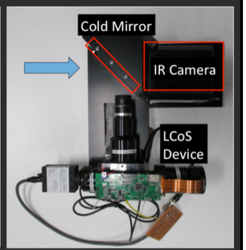
Recovering video: ping-pong ball

Reconstruction results(16 sub-frames from frame t)

Anonymous camera (Privarcy protected)



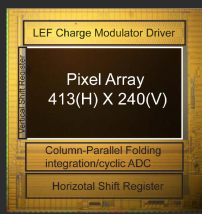
Optically masking face during surveillance



Prototype camera

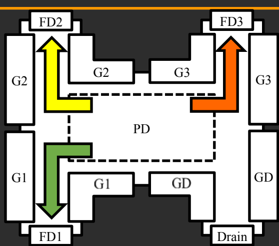
Privacy protected surveillance

Multi-tap image sensor



Han et al, Journal of the electron devices society,2014

Mechanism of capture 3 images



PD : Photo diode
FD : Floating diffusion
G1, G2, G3, GD : Gate of each floating diffusion

Exposure of Multi-tap Sensor

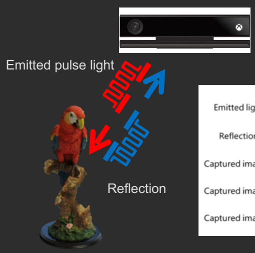
Time of flight (TOF) depth cameras

MS Kinect V2

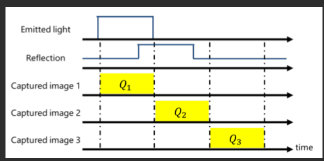
Asus Xtion2

Time-of-flight (TOF)

- Camera emits a pulse light and receives the reflected light
- Calculating the object depth from delay of the light
- The delay is calculated by ratio of the intensities of the IMGs.



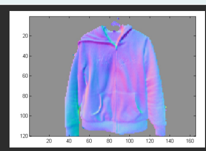
$$d = \frac{1}{2} c\Delta t \frac{Q_2 - Q_3}{Q_1 + Q_2 - 2Q_3}$$



Time of flight vs Phtometric stereo

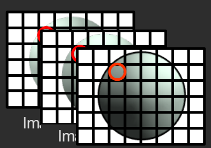
Measure distance between camera and object	Estimate the surface normal
Time-of-Flight Structured light	Photometric stereo

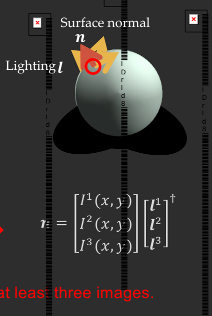

Depth map


Normal map

Photometric stereo

- The intensities of captured image depend on lighting direction and surface normal of object.



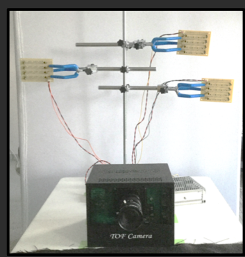


$$I^1(x,y) = \boldsymbol{n} \cdot \boldsymbol{l}^1$$

$$\boldsymbol{n} = \begin{bmatrix} I^1(x,y) \\ I^2(x,y) \\ I^3(x,y) \end{bmatrix} \begin{bmatrix} \boldsymbol{l}^1 \\ \boldsymbol{l}^2 \\ \boldsymbol{l}^3 \end{bmatrix}^{\dagger}$$

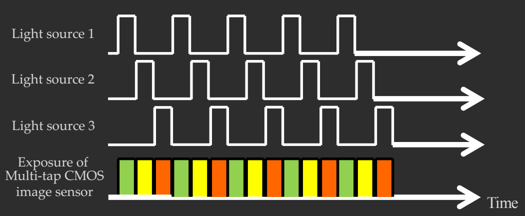
Photometric stereo requires at least three images.

Dynamic PS camera

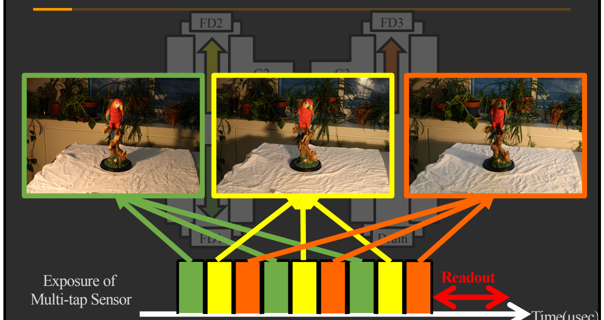


[Yoda+, ICPR2016]

Synchronization of the lights and taps

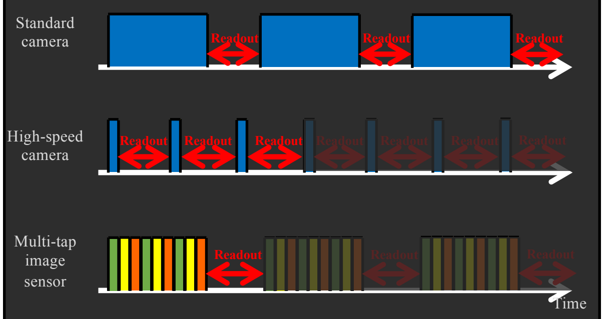


Mechanism of capture 3 images



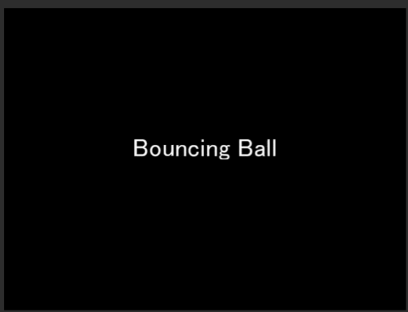
Multi-tap CMOS sensor obtains images with micro sec. difference

Exposure differences btw cameras



Comparison results

Experiments result



Conclusion

- Programmable imagers is getting popular.
- Introducing the examples of compressive video sensing and dynamic photometric stereo.
- More easy to utilize new functions of the imagers to CP and CV applications.