Space-Time-Brightness Sampling Using an Adaptive Pixel-Wise Coded Exposure

Hajime Nagahara
Osaka University
2-8, Yamadaoka, Suita, Osaka, Japan
nagahara@ids.osaka-u.ac.jp

Toshiki Sonoda
Kyushu University
744 Motooka, Nishiku, Fukuoka, Japan
sonoda@limu.ait.kyushu-u.ac.jp

Dengyu Liu
Intel Corporation
2200 Mission College Blvd Santa Clara, CA 95054
dengyu.liu@intel.com

Jinwei Gu
Nvidia Research
2788 San Tomas Expy, Santa Clara, CA 95051
jinweig@nvidia.com

Abstract

Most conventional digital video cameras face a fundamental trade-off between spatial resolution, temporal resolution and dynamic range (i.e., brightness resolution) because of a limited bandwidth for data transmission. A few recent studies have shown that with non-uniform space-time sampling, such as that implemented with pixel-wise coded exposure, one can go beyond this trade-off and achieve high efficiency for scene capture. However, in these studies, the sampling schemes were pre-defined and independent of the target scene content. In this paper, we propose an adaptive space-time-brightness sampling method to further improve the efficiency of video capture. The proposed method adaptively updates a pixel-wise coded exposure pattern using the information analyzed from previously captured frames. We built a prototype camera that enables adaptive coding of patterns online to show the feasibility of the proposed adaptive coded exposure method. Simulation and experimental results show that the adaptive space-time-brightness sampling scheme achieves more accurate video reconstruction results and high dynamic range with less computational cost, than previous method. To the best of our knowledge, our prototype is the first implementation of an adaptive pixel-wise coded exposure camera.

1. Introduction

Most conventional digital video cameras face a fundamental trade-off between spatial resolution, temporal resolution and dynamic range (*i.e.*, brightness resolution), because of a limited bandwidth for data transmission and a delay in A/ D conversion. For the trade-off between spatial resolution and temporal resolution, a few studies [5, 7, 14] have successfully used non-uniform space-time sampling (often implemented as pixel-wise coded exposure), by incorporating either smoothness in the spatial and temporal domain or sparsity in space-time volumes for reconstruction. To apply a high dynamic range (HDR) to the moving scene, Nayar and Mitsunaga [11] achieved one shot HDR imaging using a filter mosaic that has different densities on neighboring pixels. Despite their effectiveness, these sampling schemes are pre-defined, fixed, and independent of the target scene, which may be non-optimal for the recovery of long videos. For instance, static regions should be sampled at higher spatial resolution with longer exposure so as not to waste the amount of light, while moving regions should be sampled with pixel-wise coded exposure and reconstructed using a sparse representation. Nayer and Branzoi [10] captured a scene with changing their image pixel-wise exposure setting adaptively, and achieved high space-brightness resolution. However this method cannot apply to moving scene.

In this paper, motivated by these factors, and building on previous work [7], we propose an adaptive space-time-brightness sampling method to systematically optimize spatial, temporal, and brightness resolution for video capture with pixel-wise random coded exposure. This method adequately allocates the hardware resources to scene resolution within a conventional bandwidth. This is achieved by applying pixel-wise coded exposure to the moving regions and the HDR exposure coding to static regions. Our contributions include:

- **Adaptive scene sampling**. The scene content is captured with high spatial-temporal-brightness resolution by adaptively changing the pixel-wise coded exposure patterns as a feedback loop. Conventional methods can tackle only one of these trade-offs with a fixed and predefined sampling scheme, or using expensive hardware that has a large bandwidth for any capturing frame.

- **High space-time-brightness resolution**. This is equivalent to "motion-aware" sampling. Note that it is different to flexible voxels [5] that have fixed space-time sampling and motion-aware reconstruction. The moving region applied to random code for improving space-time resolution and the static region is applied to HDR code for increasing brightness resolution. For motion detection, we used simple inter-frame subtraction, although other existing motion detection methods can also be used.

We performed simulations for validation. The simulations were performed with real video sequences captured by high-speed cameras that have high brightness resolution (16 bits) as the ground truth. Frames from the videos were used to synthesize the coded images captured by pixel-wise coded exposure. These images have similar characteristics to the coded images captured by a real sensor. The simulation results are shown in Section 4.1 and Fig.4.

We also built a prototype camera with adaptive pixel-wise coded exposure, and carried out real experiments to show the feasibility of adaptive exposure coding in practice. While being intuitive, implementing the above ideas in hardware is nontrivial because there is no commercial image sensor that can use a pixel-wise exposure. Thus, we must demonstrate the effectiveness of our proposed method using other optical devices. Most previous work [7, 14] has used a spatial light modulator (SLM) such as a digital micromirror device (DMD) or liquid crystal on silicon (LCoS). These SLMs often can only update preloaded patterns on the fly. We also used a LCoS which can adaptively display patterns via DVI video interface. A PC generates the adaptive coding patterns and display the patterns to the LCoS from the feedback of the former captured image. We built the prototype to capture the adaptive coded exposure video in real time for real experiments.

The real experiment results shown in Section 4.2, Fig. 6 demonstrate the effectiveness of our proposed adaptive pixel-wise coded exposure.

2. Related Work

Scene adaptive sampling and reconstruction. Nayar and Branzoi [10] adaptively changed the throughput of the incoming light by pixels using a liquid crystal display (LCD) and achieved HDR imaging. They achieved high spatial and brightness resolution, but not temporal resolution. They adaptively change the density of pixels form the feedback of previous frame so that the pixel avoid a saturation of the brightness range. There are a few studies of adaptive changing to capture or reconstruct a video. Yang *et al.* [22] proposed to adaptively change the number of Gaussian mixture model (GMM) basis for compressive video reconstruction. Yuan *et al.* [24] adaptively changed the temporal compression rate based on the velocity of the motion. Warnell *et al.* [20] proposed to adaptively change the number of measurements for background subtraction.

Efficient video capture for high spatial-temporal resolution. There are two approaches to overcome the fundamental trade-off between spatial and temporal resolution for video capture. (1) With multiple cameras, multiple video sequences can be combined to obtain complementary information [16, 4, 21, 1]. (2) With a single camera, prior studies have focused on the design of the shutter function (*i.e.*, space-time sampling schemes) and the reconstruction with prior information (*i.e.*, sparsity, smoothness, motion, *etc.*). Examples of non-uniform space-time sampling schemes include flutter shutter [8], flutter shutter for periodic motion [18], coded rolling shutter [3], hybrid grid shutter [2, 5], and pixel-wise coded exposure [7, 14]. For reconstruction with prior information, the smoothness in either spatial or temporal domains is used for motion-ware interpolation [2, 5]. Sparsity has been extensively used [19, 15, 17, 7], as well as other constraints such as optical flow [14]. Despite their effectiveness, these methods use predefined space-time sampling schemes that are fixed over time. These sampling schemes are also independent of the scene content, which is good for the recovery of a single coded image but may be non-optimal for the recovery of multiple consecutive coded images. In contrast, our method uses an adaptive approach that updates the space-time sampling scheme for efficient video capture.

Our method is related to the work of Lichtsteiner *et al.* [9], who built a new image sensor to detect and capture only moving regions for output. However, their method has limited spatial resolution and requires a specially designed image sensor.

High dynamic range imaging. To use HDR imaging with a normal commercial camera (many cameras have only 8-bit brightness resolution), one captured multiple images of the same scene with different exposures and combined them. Because this technique is prone to errors when there is motion in the scene or camera, two types of approach have been studied for such a moving scene ($i.e.$, high temporal resolution). First, to compensate of the difference between multiple captured images whose appearance slightly changed, and then analyzing them with post-processing [6] ($e.g.$, optical flow) or removing the motion blur [23].

Second, use special hardware. While a normal camera uniformly samples scene intensity by all pixels, Nayar and Mitsunaga [11] placed a density mosaic filter on their im-

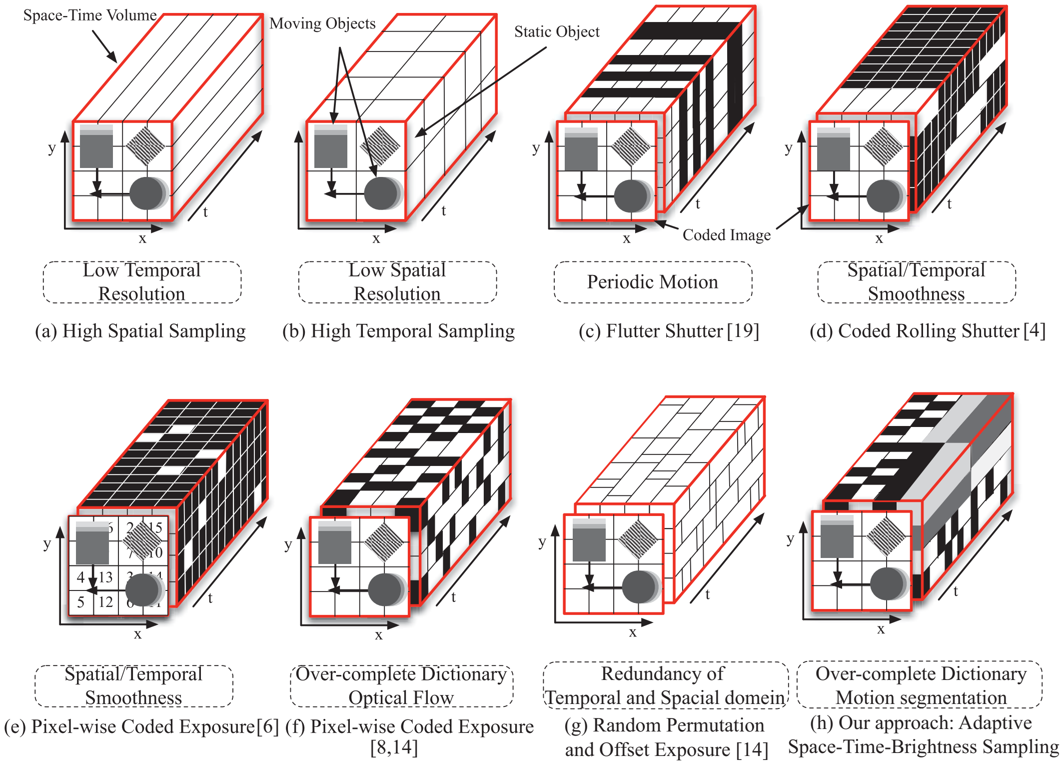


Figure 1. Overview of our work and related space-time sampling schemes. When capturing a space-time volume (red rectangular box), conventional digital cameras can either have (a) dense spatial sampling with coarse temporal sampling or (b) vice-versa . (c) By strobing the exposure, the flutter shutter is used to recover periodic motion. (d) Coded rolling shutter is proposed to control the readout timing and exposure length for each row of CMOS sensors. (e) A mixture of denser spatial samples and temporal samples are implemented as a grid shutter for motion-aware high-speed imaging. (f) Pixel-wise coded exposure has been recently implemented for efficient video capture. (g) Several different exposure offsets are randomly arranged on the spatial–temporal volume. There is no blocking between the exposure times, and no wasted light. A variety of priors and constraints (dashed line boxes in (c)–(g)) are exploited for video reconstruction from a few coded images (red square boxes). Nevertheless, in these work, both the coded exposure pattern and the priors are fixed. In our approach (h), we adaptively change the coded exposure patterns (*e.g.*, pixels in moving regions are randomly exposed for space-time recovery and pixels in the static diamond are exposed for HDR).

age sensor and made the sensor have spatially different exposures. They successfully obtained information with one shot that was equivalent to the information of several shots. However, this method degrades the original spatial resolution of the image sensor, because a pixel that has high brightness resolution is constructed from four pixels that have low brightness resolution. This sampling scheme is similar to that of a Bayer pattern when capturing a color image (the HDR mosaic [11] samples the intensity not the spectrum). It is difficult to change how to sample the scene adaptively because the densities of the filter is optically fixed.

A few studies have attempted to simultaneously achieve efficient video capture and high dynamic imaging. As we have shown above, Gu *et al.* [3] also developed a method to reconstruct a video from the coded captured image. The image recorded the information of motion and its high brightness resolution on a 2D plane using a coded rolling shutter, but it results in degradation of the spatial resolution. The purpose of the study of Portz *et al.* [13] is most similar to this study. They used several different exposure offsets randomly arranged on the spatial-temporal volume and attempted to reconstruct the video whose space-time-brightness resolutions are all high. That video was reconstructed by exploiting the redundancy of the spatial and temporal volume. This method also repeatedly used the same fixed sampling pattern that was pre-defined and independent of the scene content, and it only showed the feasibility without any real experiments.

3. Adaptive Pixel-wise Coded Exposure

We propose space-time-brightness sampling by pixel-wise coded exposure. We adaptively switch the coded patterns, based on the motions of a scene: the pixels in the moving region are randomly exposed and those in the static region are exposed for the HDR, as shown in Fig. 1(h).

Table 1. Comparison of Space-Time Sampling Schemes

Method	Sampling Function	Reconstruction	Hardware	Limitation
Wakin *et al.* [19]	Pixel-wise Random $S(x, y, t)$	Greedy Algorithm Sparsity Constraint	DMD Beam Splitter	Not suitable for video
Veeraraghvan *et al.* [18]	Flutter Shutter $S(t)$	l_1-norm Minimization	Ferroelectric Shutter	Only for periodical video
Gu *et al.* [3]	Coded Rolling Shutter $S(y, t)$	Interpolation Optical Flow	CMOS sensor with modified control unit	Lack flexibility on vertical direction
Gupta *et al.* [5]	Pixel-wise Grid $S(x, y, t)$	Interpolation Optical Flow	Projector Beam Splitter	Ambient illumination, low SNR
Reddy *et al.* [14]	Pixel-wise Random $S(x, y, t)$	Sparsity Constraint Optical Flow	LCoS Beam Splitter	Multiple coded images required
Hitomi *et al.* [7]	Pixel-wise Random $S(x, y, t)$	Greedy Algorithm Dictionary Learning	LCoS Beam Splitter	background low SNR, non-adaptive dictionary
Portz *et al.* [13]	Random permutation and offset of different exposures $S(x, y, t)$	Exhaustive search for the K-nearest space-time patches	No real experiment	No real experiment

Figure 1 summarizes several space-time sampling schemes. Assume we capture a space-time volume (the red rectangular box) with high-speed motion objects ($e.g.$, the moving square and circle) and high-texture static objects ($e.g.$, the diamond). With a limited bandwidth, a high spatial resolution camera can capture the texture on the static object, but this results in motion blur of the moving object. In contrast, a camera with a high temporal resolution can capture the motion but fails to preserve the texture. For conventional digital video cameras, the space-time sampling is constant, $i.e.$, $S(x, y, t) = 1$. Fig. 1(c)–(g) show recent flexible space-time sampling schemes that aim to exploit redundancy in videos for efficient video capture. Flutter shutter [18] is a 1-D function $S(t)$ used for the recovery of periodic motion. Coded rolling shutter [3] controls the readout timing and the exposure length in a row-wise manner, which is a 2-D function $S(y, t)$. Recently, full 3-D sampling $S(x, y, t)$ as pixel-wise coded exposure has been implemented [5, 7, 14] and incorporated in a variety of priors and constraints for the reconstruction, including spatial/temporal smoothness, optical flow, and sparsity. Portz *et al.* [13] only validated their method using simulation experiments, and they achieved efficient spatial-temporal-brightness sampling with random per-pixel exposure times and offsets. Table 1 compares these methods in more detail. Nevertheless, both the sampling schemes and representations are fixed over time.

We aim to develop an adaptive sampling scheme for the recovery of long videos. Figure 2 shows an overview of the process. Here, we define the frame f as the unit of the captured image and time t indicates the latent high temporal images inside the captured frame. We generate a motion segmented mask from last two capturing images. We adaptively change different exposure coding patterns based on the motions of the regions in the mask. We apply a random exposure pattern for the moving regions, and a density mosaic pattern for the static region from the real-time feedback of the segmentation result. We also separately reconstruct the images based on the segments. The moving regions of the images are estimated by compressive video reconstruction. The static regions of the images are generated by HDR image estimation. These regions of the images are integrated to the output image as

$$I_t^{output} = \hat{E}_t \bigcup I^{HDR}, \qquad (1)$$
$$\hat{E}_t \bigcap I^{HDR} = \emptyset,$$

where I_t^{output} is an output image that has a high spatial-temporal-brightness resolution, and $\hat{E}_t$ and I^{HDR} are the reconstructed moving regions and the static region of the images, respectively.

The benefits of our proposed methods are twofold:

- Adaptively pixel-wise exposure based on motion. Dynamic region and static region are adaptively applied to random and HDR code of exposures respective to motion segmentation.
- Enhancing space-time resolution for moving region and brightness resolution for static region. Also by only applying compressive video reconstruction to moving region, we can reduce computational cost to the previous approaches.

In the latter part of this section, we describe the detailed coding and reconstruction methods for moving and static regions in Section 3.2 and Section 3.3, respectively.

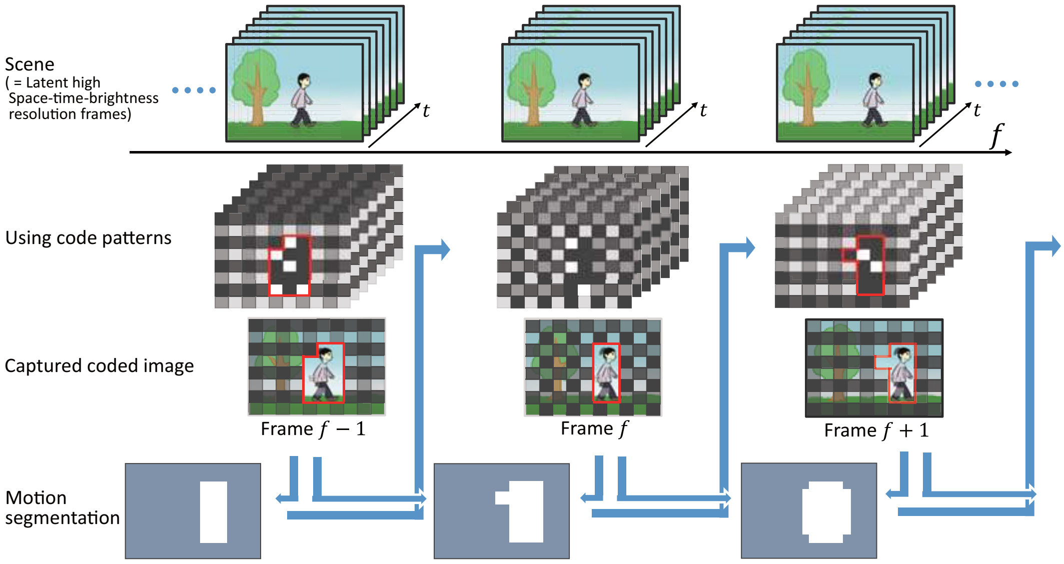


Figure 2. Overview of process for generating our adaptive coded exposure. It shows how to generate the exposure pattern to code frame $f+1$ after we obtain frame f. **The top row** is the real scene that has high spatial-temporal-brightness resolution. **The second row** are the generated exposure patterns used. **The third row** are the captured coded images. **The bottom row** is a workflow of motion segmentation. Firstly, we subtract the current frame f from previous frame $f-1$ and segment the moving/static region. Before subtraction, coded exposure must be compensated with the corresponding known spatially varying exposure pattern. According to the obtained segmentation, the new coding exposure pattern is generated. The region corresponding to the moving region consists of the random exposure patterns, and the region corresponding to the static region consists of a HDR exposure pattern. It is then applied to code the next capturing scene.

3.1. Motion segmentation for adaptive coding

We propose to adaptively choice the exposure code patterns region by region in a capturing frame. We assume that dynamic regions are changing region caused by object motions and camera motion etc. We use simple inter-frame subtraction between last two frames $f-2$ and $f-1$ to generate the motion mask at frame f, as shown in Fig. 2. We get the difference image from the subtraction and apply thresholding and dilation to obtain the motion segmentation mask for frame f. Random exposure coding and HDR coding are applied to the dynamic and static regions respectively based on the motion mask. We repeat this process for all the frames to achieve adaptive coding.

3.2. Space-time coding and reconstruction for moving region

Our work is based on Hitomi's method [7] for the moving regions. In the following, we give a brief summary of the method [7].

Let $E(x, y, t)$ denote the target video and $I(x, y)$ be the captured coded exposure image, we then have

$$I(x,y) = \sum_{t=1}^{N} S(x,y,t) \cdot E(x,y,t), \tag{2}$$

where N is the number of frames within the target volume.

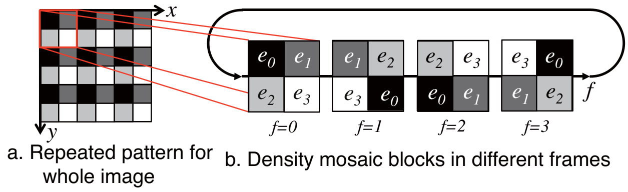


Figure 3. Coded exposure pattern for spatial brightness sampling.

Specifically, each voxel in $E(x, y, t)$ is assumed to be a sparse linear combination of some basis motion patterns from a learned overcomplete dictionary $\mathbf{D} = [D_1(x,y,t), D_2(x,y,t), \cdots, D_K(x,y,t)]$, *i.e.*, $\mathbf{E} = \mathbf{D}\boldsymbol{\alpha}$. Equation (2) can be rewritten in a matrix form as

$$\mathbf{I} = \mathbf{S}\mathbf{E} = \mathbf{S}\mathbf{D}\boldsymbol{\alpha}. \tag{3}$$

The over-complete dictionary $\mathbf{D}$ is learned from a random collection of videos. $\mathbf{D}$, $\mathbf{S}$, $\mathbf{I}$, and $\hat{\boldsymbol{\alpha}}$ are estimated using standard sparse reconstruction techniques such as orthogonal matching pursuit (OMP) [12], *i.e.*,

$$\min \parallel \alpha \parallel_0 \; s.t. \parallel \mathbf{I} - \mathbf{S}\mathbf{D}\boldsymbol{\alpha} \parallel_2^2 \leq \varepsilon, \tag{4}$$

and $\hat{\mathbf{E}}$ is represented as $\hat{\mathbf{E}} = \mathbf{D}\hat{\boldsymbol{\alpha}}$.

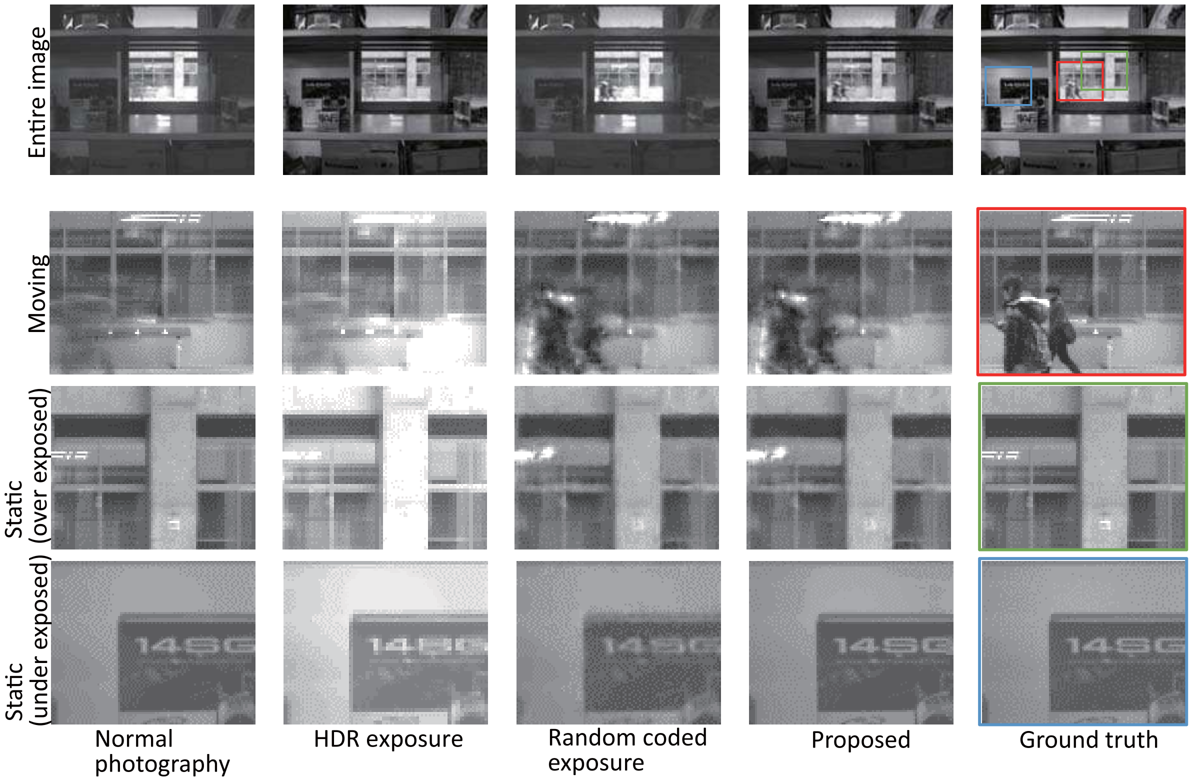


Figure 4. Results of a simulation experiment with some other methods for comparison. For ground truth video, we captured outdoor scene from inside of the room using a high-speed camera with high brightness resolution. One of the captured frames is shown in the right column. The other column images are generated from ground truth to imitate the images obtained using normal photography, HDR exposure [11], random exposure [7], and our proposed method. The entire images that has high dynamic range(*i.e.*, Spatial varying exposure, our proposed and Ground truth) are shown with tonemapped. For easy to see and fair comparison, each row of zoomed images are adjusted with the same tonecurve. See the reconstructed video data in the supplementary material.

3.3. Spatial brightness coding and reconstruction for static regions

We apply high dynamic imaging using spatially varying exposure to the static regions of a scene. Similar to the HDR mosaic pattern [11], we also use the mosaic of four different densities with every four neighbor pixels, as shown in Fig. 3. Fig. 3b shows the zoom up portion of a unit of four neighboring patterns as a mosaic block. The densities of the pattern make the different sensitivities or exposures e_i, where $e_0 < e_1 < e_2 < e_3$. The patterns of the block are recursively changed by the frame f, as shown in Fig. 3b, and each mosaic block is repeated over all of the static regions of the image, as shown in Fig. 3a. We formulate the space-time exposure pattern as

$$S(x, y, f) = e_{(2y+x+f) mod 4}. \quad (5)$$

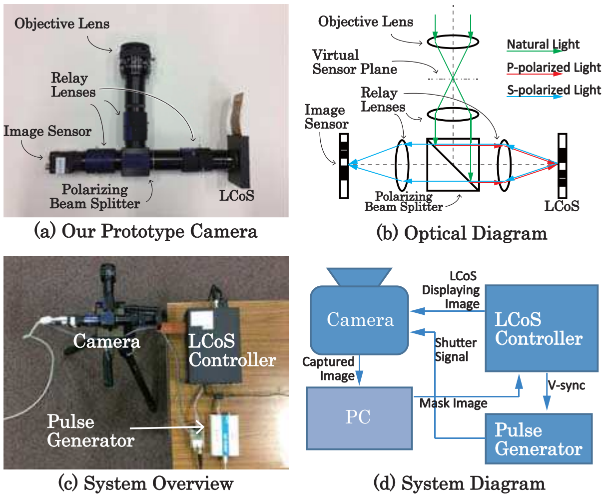
We also describe the captured image with the coded exposures as

$$I(x, y, f) = S(x, y, f) \cdot E(x, y). \quad (6)$$

After we obtain four consecutive frames, we can simply reconstruct the HDR image at the frame f by

$$I^{HDR}(x, y, f) = \sum_{i=0}^{3} I(x, y, f - i)/S(x, y, f - i). \quad (7)$$

If we cannot obtain four continuous full frames as the static region, we interpolate the lacking exposures of the pixel from the neighboring pixels of the exposure. For this, we use the nearest neighbor interpolation method in our experiments. The proposed HDR exposure pattern is similar to Nayar's HDR pattern [11]. However, we also use temporal changes of the patterns and reconstruct the original spatial resolution, while Nayar's HDR pattern is temporally constant and the spatial resolution decreases by one-quarter .



(a) Our Prototype Camera (b) Optical Diagram

(c) System Overview (d) System Diagram

Figure 5. A prototype of our adaptive coded exposure camera system. (a) and (b) show the overview of the prototype camera and its optical diagram. (c) and (d) show the overview of the entire system and a diagram of signal connections between the camera and the other equipment.

4. Experimental Results

4.1. Simulation

The simulation results for adaptive coded exposure are shown in Fig. 4. We obtained the ground truth video using a high-speed camera (Point Grey GS3-U3-23S6C) with high brightness resolution: spatial resolution 480 $\times$ 400, temporal resolution 180 fps, and brightness resolution 16 bit . We compare our adaptive sampling scheme with normal photography (low temporal and brightness resolution), HDR exosure [11] (low temporal resolution and high brightness resolution), and random exposure [7] (high temporal resolution and low brightness resolution). In Fig. 4, the top rows shows one of the complete images from the video, (the images of [11], [7], and our image are reconstructed images). The other rows are the zoomed up one according to the properties. Our proposed procedure works well and obtained good image quality in all of the zoomed up regions compared with the conventional methods. Thus, our proposed method can sample the scene information adaptively and correctly.

4.2. Real Experiment

We built a prototype coded exposure camera to show the feasibility of our proposed motion-adaptive coded exposure method. Fig. 5 shows the overview of the prototype camera. It consists of an object lens (Tokina f = 12.5 mm), three relay lenses, a polarizing beam splitter, LCoS (Holoeye LC-R720), and CCD camera (Point Grey GS3-U3-28S5M). The LCoS and CCD were connected to a PC (Core i7, 3.3 GHz) via a DVI and USB3.0 interface, respectively. The refresh rate of the LCoS was 180 Hz and the patterns were adaptively given by the DVI video interface from the PC. The pulse generator generated the CCD shutter signal from the V-sync of the LCoS display. The CCD was completely synchronized by generating a 1:36 ratio of the V-sync. A coded video was captured at 5 fps and each frame was coded by 36 exposure patterns, which was the same as the simulation experiments. The PC adaptively generated the adaptive coded exposure pattern in real time. Thus, we generated a 180 fps video after the reconstruction. We calibrated the corresponding pixels between the LCoS and CCD, and picked the centered pixels of the CCD to make the coded exposure image, because the pixel size of the LCoS was three times larger than the CCD pixel size. Finally, we obtained 500$\times$600 pixels of the coded video.

Fig. 6 shows the result of the real experiment: the captured images, adaptive moving/static segmentation and reconstructed video frames. The top row of the figure shows three captured frames from the prototype camera and the other rows of the figure show some patterns (t = 15, 30) of 36 moving/static segmentations and the corresponded reconstruction images as a page limitation. The segmented moving region of the walking man slightly moved between the three captured frames, because the masks were adaptively generated by the motion of the previous frame. The captured images were coded by adaptive pixel-wise exposure, the moving region of the scene was randomly sampled, and static region was sampled with spatially varying exposure. The third row of the figure shows some of the reconstructed frames at 180 fps (= 5 fps $\times$ 36 coded patterns). The man is walking in the reconstructed video. The detail of outside the room can be seen, while the original target scene has wide dynamic range and the exposure is set for inside the room. Thus, we showed that adaptive coded exposure can also work online with the prototype camera.

5. Conclusions and Discussions

In this paper, we propose an efficient way to capture video by adaptive pixel-wise coded exposure. According to the scene content, an efficient sampling scheme is automatically selected. Random exposure is only applied to moving regions in the video to reduce reconstruction time. For static regions (*e.g.*, the background), HDR exposure is used to obtain high brightness information. We demonstrated the quality of the reconstructed video by simulation. In addition, we built a prototype camera and showed the feasibility of the real-time adaptive coding in real experiments.

Our approach and current implementation have a few limitations. The effectiveness strongly depends on the accuracy of the moving/static region segmentation. While the main aim was to propose an adaptive sampling scheme, for improvement of this method, it is necessary to consider the

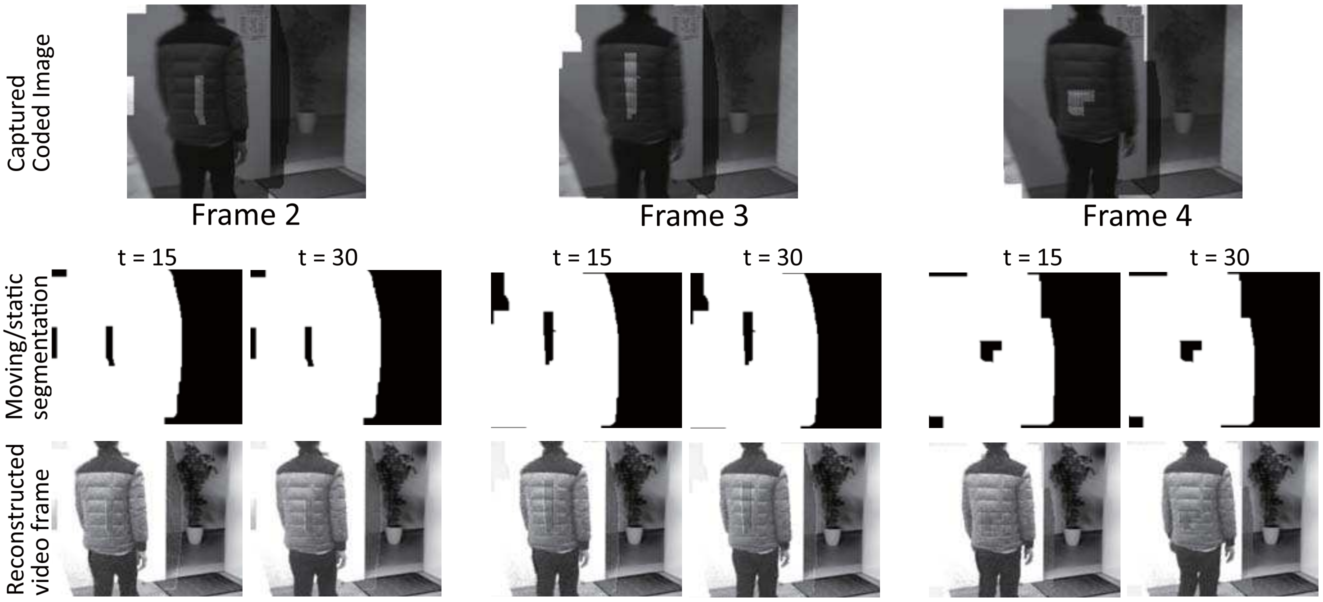


Figure 6. Results of real experiments. Three consecutive frames (Frame 2 – 7) are extracted from the captured coded video. **Top row** shows captured coded images. **Second row** shows the moving/static region segmentations. Note that 36 patterns were used to code each captured frame, but here we only show two patterns (t = 15, 30). They were generated from image analysis of the former frame, so they change each frames of the captured video. **Third row** shows the reconstructed and tonemapped images from the coded images with the ratio of $\times 36$. See the reconstructed video data in the supplementary material.

use of more sophisticated segmentation. In our current implementation, there is three frames latency between the motion detection and capturing with the adaptive pattern for the reagion. The motion blur or satulation would be appeared in the first frame when the object or the motion is suddonly appeared like a commertial adaptive exposure camera. We ideally need a special CMOS imager which can detect the motion and apply the adaptive exposure on chip for eliminating the latency.

References

[1] A. Agrawal, M. Gupta, A. Veeraraghavan, and S. Narasimhan. Optimal Coded Sampling for Temporal Super-Resolution. In *CVPR*, pages 599–606, 2010.

[2] G. Bub, M. Tecza, M. Helmes, P. Lee, and P. Kohl. Temporal Pixel Multiplexing for Simultaneous High-Speed, High-Resolution Imaging. *Nature Methods*, 7, 2010.

[3] J. Gu, Y. Hitomi, T. Mitsunaga, and S. Nayar. Coded Rolling Shutter Photography: Flexible Space-Time Sampling. In *ICCP*, pages 1–8, 2010.

[4] A. Gupta, P. Bhat, M. Dontcheva, O. Deussen, B. Curless, and M. Cohen. Enhancing and Experiencing Space-Time Resolution with Videos and Stills. In *ICCP*, pages 1–9, 2009.

[5] M. Gupta, A. Agrawal, and A. Veeraraghavan. Flexible Voxels for Motion-Aware Videography. In *ECCV*, volume 3, page 6, 2010.

[6] M. Gupta, D. Iso, and S. Nayar. Fibonacci Exposure Bracketing for High Dynamic Range Imaging. In *ICCV*, pages 1–8, Dec 2013.

[7] Y. Hitomi, J. Gu, M. Gupta, T. Mitsunaga, and S. Nayar. Video from a Single Coded Exposure Photograph using a Learned Over-Complete Dictionary. In *ICCV*, pages 287–294, 2011.

[8] J. Holloway, A. C. Sankaranarayanan, A. Veeraraghavan, and S. Tambe. Flutter Shutter Video Camera for Compressive Sensing of Videos. In *ICCP*, pages 1–9. IEEE, 2012.

[9] P. Lichtsteiner, C. Posch, and T. Delbruck. A 128 x 128 120 dB 15 us Latency Asynchronous Temporal Contrast Vision Sensor. *IEEE Journal of Solid-State Circuits*, 43(2), 2008.

[10] S. Nayar and V. Branzoi. Adaptive Dynamic Range Imaging: Optical Control of Pixel Exposures over Space and Time. In *ICCV*, volume 2, pages 1168–1175, 2003.

[11] S. Nayar and T. Mitsunaga. High Dynamic Range Imaging: Spatially Varying Pixel Exposures. In *CVPR*, volume 1, pages 472–479, Jun 2000.

[12] Y. Pati, R. Rezaiifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, pages 40–44 vol.1, 1993.

[13] T. Portz, L. Zhang, and H. Jiang. Random coded sampling for high-speed HDR video. In *ICCP*, Boston, MA, April 2013.

[14] D. Reddy, A. Veeraraghavan, and R. Chellappa. P2C2: Programmable Pixel Compressive Camera for High Speed Imaging. In *CVPR*, pages 329–336, 2011.

[15] A. C. Sankaranarayanan, C. Studer, and R. G. Baraniuk. CS-MUVI: Video Compressive Sensing for Spatial-Multiplexing Cameras. In *ICCP*, pages 1–10, 2012.

[16] E. Shechtman, Y. Caspi, and M. Irani. Space-time super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):531 –545, april 2005.

[17] X. Shu and N. Ahuja. Imaging via three-dimensional compressive sampling (3DCS). In *ICCV*, pages 439–446, 2011.

[18] A. Veeraraghavan, D. Reddy, and R. Raskar. Coded Strobing Photography: Compressive Sensing of High Speed Periodic Videos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4):671–686, 2011.

[19] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R. Baraniuk. Compressive Imaging for Video Representation and Coding. In *Picture Coding Symposium*, 2006.

[20] G. Warnell, S. Bhattacharya, R. Chellappa, and T. Basar. Adaptive-rate compressive sensing via side information. *IEEE Transactions on Image Processing (TIP)*, 24:3846–3857, 2015.

[21] B. Wilburn, N. Joshi, V. Vaish, M. Levoy, and M. Horowitz. High-Speed Videography using a Dense Camera Array. In *CVPR*, volume 2, pages 294–301, 2004.

[22] J. Yang, X. Yuan, X. Liao, P. Llull, D. J. Brady, G. Sapiro, and L. Carin. Video compressive sensing using gaussian mixture models. *IEEE Transactions on Image Processing (TIP)*, 23:4863–4878, 2013.

[23] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. In *ACM SIGGRAPH 2007 Papers*, SIGGRAPH ’07, New York, NY, USA, 2007. ACM.

[24] X. Yuan, P. Llull, X. Liao, J. Yang, G. Sapiro, D. J. Brady, and L. Carin. Low-cost compressive sensing for color video and depth. In *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2014.