# A Flexible Programmable Camera Control and Data Acquisition Hardware Platform

Fei Cheng[1]([✉]), Jimin Xiao[1], Tammam Tillo[1], and Yao Zhao[2]

[1] Department of Electrical and Electronic Engineering,
Xian Jiaotong-Liverpool University (XJTLU), 111 Ren Ai Road, SIP,
Suzhou 215123, Jiangsu Province, People's Republic of China
{fei.cheng,jimin.xiao,tammam.tillo}@xjtlu.edu.cn
http://www.mmtlab.com
[2] Beijing Jiaotong University, Institute of Information Science, Beijing, China
yzhao@bjtu.edu.cn

**Abstract.** There are a number of standard video sequences produced by some organizations and companies, which simulate different scenes and conditions in order to properly test video coding methods. With the rapid development of sensing and imaging technologies, more information can be employed to assist video coding. However, the existed sequences cannot satisfy the recently proposed video coding methods. Moreover, to capture 3D video, multi-cameras, including texture and depth cameras, are used together. Thus, a synchronization mechanism is needed to control different cameras. In addition, some adverse conditions, such as light source flicker, affect the quality of the video. In order to acquire proper video sequences, we proposed a flexible control platform for texture and depth cameras, which could address the above issues. We have applied the platform to test several new methods proposed recently. With the assistance of this platform, we have achieved considerable positive results.

**Keywords:** Video sequence · Camera control · Synchronization · Flicker · Platform

## 1 Introduction

Globally, video traffic accounted for 66 percent of all consumer Internet traffic in 2013 [1]. With the development of imaging and display technologies, the video resolution is rapidly increasing. The 4K ($3840 \times 2160$) resolution has been used in camcorders and TVs. Meanwhile, 3D TVs and films are being accepted in consumer market. Moreover, some sensors, such as motion sensors and GPS, have been integrated on smart devices, which could assist the video coding and enhance the video quality [2]. Figure 1 shows the current development trend of
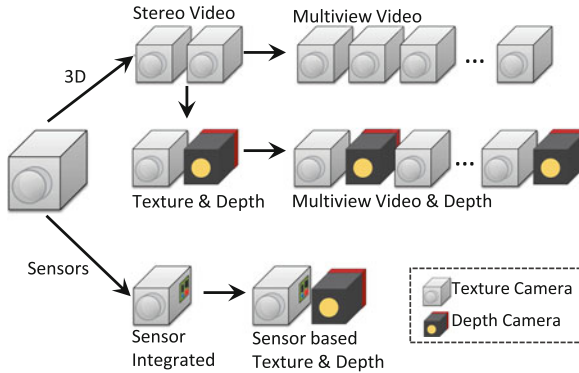
**Fig. 1.** The development trend of video acquisition

video coding. The compression performance of modern video coding is improving constantly, and new video coding methods are emerging. Therefore, new video sequences with auxiliary information need to be produced to test new methods, and more importantly video acquisition platforms with sensor capability are needed to simulate different emerging scenarios.

In fact, the existing video sequences cannot meet the requirement of some recently proposed video coding methods. For instance, the movement of smart phone while capturing a video results in the global motion between frames, which increases the bit-rate and the computational complexity of the video encoder. The global motion information could be used to assist video coding [3]. The problem is that no existing video sequence includes the camera global motion information. Therefore, it is necessary to develop an open and flexible video capturing platform to obtain different kinds of video and auxiliary information.

In terms of 3D video, complicated camera settings bring a great challenge for 3D video acquisition. Multiple cameras have to work synchronously to capture the 3D scene. In stereo video, two cameras are employed to capture two views, while the multiview video normally deals with cases with more than two cameras. Another 3D video scheme is texture plus depth format, which generates virtual views using the Depth-Image-Based Rendering (DIBR) [4,5]. A depth map, which represents the distance from the objects in the scene to the camera, and its aligned texture, have been exploited to describe 3D scenes. Different approaches to obtain depth exist, however, currently most depth maps are estimated using software-based approaches from several views of the scene, which has many problems. Meanwhile, the hardware depth acquisition technology is rapidly developing. Some depth cameras are now becoming available. Multi-view Video plus Depth (MVD) format is a promising way to represent 3D scene [6,7], where multiple texture and depth cameras are employed. As the alignment of texture and depth is of great importance in 3D video, the position of each camera and the frame rate need to be controlled accurately.

There are some other issues affecting the video sequence acquisition. Firstly, the flicker effect caused by light source makes the brightness fluctuation between frames [8]. This effect would reduce the performance of Motion Estimation (ME) for video encoder [9]. This problem becomes more serious with different types of cameras, which causes luminance difference between different views. Secondly, without a hardware trigger, it is difficult to obtain several video acquisitions for the same movement. Lastly, the existing camera platforms are not flexible and programmable to fully control cameras and their movements.
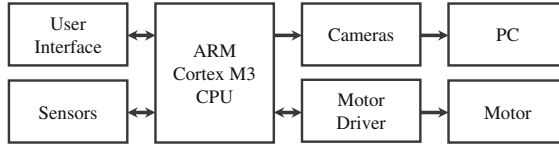


**Fig. 2.** The diagram of the proposed platform

We proposed and developed a programmable camera control and data acquisition hardware platform as shown in Fig. 2. This platform is designed based on a slider tracker, while a DC motor can move the platform properly. An ARM CPU is exploited to control and manage the motor, sensors and interface. This platform can drive four texture cameras or depth cameras by using Pulse Width Modulation (PWM). To reduce the flicker effect, the frequency and phase of power grid can be sampled. Moreover, in order to fully and conveniently control cameras and related devices, a user-friendly interface is developed. This platform has been used to obtain experiment data for several schemes we proposed recently. We have achieved some positive results based on those video sequences and data. Two works about planar surfaces detection have been published [10]. Figure 3 presents an example of this platform, which is able to record texture plus depth video with global camera motion.
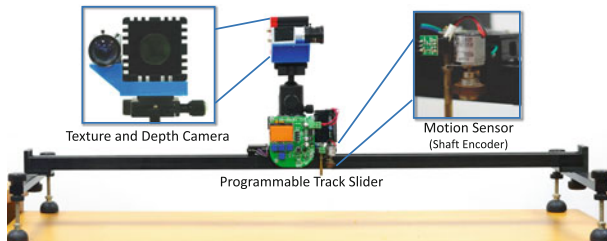


**Fig. 3.** An example of our texture plus depth record system with global motion information

The rest of this paper is organized as follows. In Sect. 2, the details of our solutions and designs are introduced. Then, tests and related experiment results are presented in Sect. 3. Finally, Sect. 4 concludes this work.

## 2   Methodology

The proposed platform is developed to satisfy the requirements of several experiments. In this section, the difficulties are specified and the corresponding solutions are explained. Then, the details of developed hardware and software are introduced.

### 2.1   Difficulties and Solutions

For texture plus depth format or multiview video sequence acquisition, a mechanism should be used to mount two different types of cameras. Their relative positions are pivotal arguments for image alignment and rendering. Although the calibration methods could reduce of position errors, the accurate positions could minimize the errors of image alignment and rendering. As different cameras have different sizes and shapes, it is difficult to mount them properly. In order to make the camera positions relatively accurate, the 3D molding and printing technology is exploited to customize the cameras holder. By using 3D molding, the holder could be designed to adjust to camera size and shape. More importantly, it could guarantee that the center of each camera is aligned horizontally or vertically.
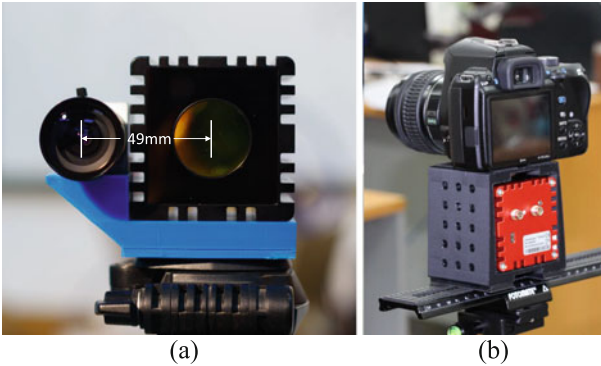


(a)                                   (b)

**Fig. 4.** 3D printed holders for different depth and texture cameras

Figure 4 displays two examples of 3D printed holders. In Fig. 4 (a), an industry camera Blaser acA640-90gc is used as texture camera, while the depth camera is a SwissRanger SR4000 [11]. The Fig. 4 (b) is a vertically alignment of a DSLR camera and depth camera.

The flicker effect is caused by unstable light sources. As the power grid is AC power, the lamp luminance would change with the frequency of power grid. A kind of common light source is fluorescent lamp, which would cause serious flicker effect. Even incandescent produced light suffers from flicker effect. As shown in Fig. 5, when the frame rate of video capturing dose not exactly match the frequency of the power grid, the video frame would flicker. Therefore, we sample the frequency and the phase of the power grid in order to adjust the frame rate and trigger the acquisition. Consequently, the video flicker effect can be minimized.
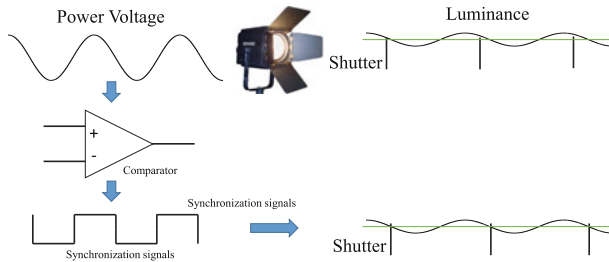


**Fig. 5.** The flicker effect and the proposed solution

In some video experiments, the video is captured when the camera is moving. If we want to obtain the same video content with different camera settings, the video capturing should start at a fixed position. We use a mechanical switch as shown in Fig. 6 to generate a trigger signal. In Fig. 6, the longer bar represents the slider tracker, while the shorter one is a trigger block. When the platform is moving towards left, the switch on the platform will hit the trigger block and generate a signal. This trigger signal is sent to the CPU, which starts the capturing process. This solution is low cost and reliable.
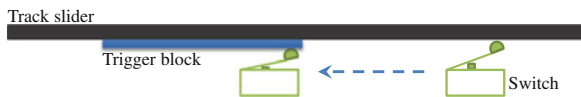


**Fig. 6.** The proposed position trigger using a mechanical switch

## 2.2   Implementation

In order to increase the reliability and flexibility, a Printed Circuit Board (PCB) was developed to integrate all the components and circuits. The PCB schematic consists of a CPU, a power supply, sensors, communication modules, camera drivers, a motor driver and an user interface, etc.

The selected CPU is an ARM 32-bit cortex M3 core based STM32 F103. The supported frequency is up to 72 MHz, which is enough for the developed control strategy. This CPU supports seven timers, which could generate PWM signals to drive motor and cameras. The power is supplied by a 12V DC battery to avoid interference from AC power. To minimize the flicker effect, a comparator is used to convert the AC signal to digital synchronization signal as shown in Fig. 5. A shaft encoder is exploited as motion sensor to adjust the speed of motor and record the position of the platform. In terms of communication, BlueTooth and RS232 are employed to communicate with smart devices and PCs. To drive different kinds of cameras, a inferred (IR) emitter and four PWM signal sockets are used. The IR emitter can control most DSLR cameras, while the PWM signal can drive industry cameras. The motor driver is composed of two half bridge ICs. In order to operate the platform conveniently, the user interface consists of an IR remote controller, buttons and a LCD screen.

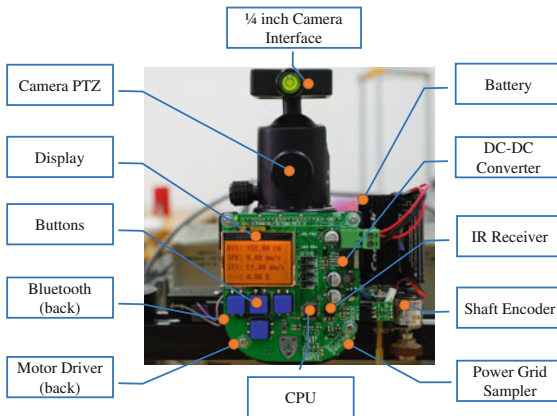Figure 7 shows the platform and each key component on it. The PCB schematic and layout are available for download at http://www.mmtlab.com/platform[1].



**Fig. 7.** The hardware design for this platform

The software of the proposed platform is developed based on the standard peripheral library published by STMicroelectronics. By using the library, the development cycle can be reduced.

The system resources need to be allocated reasonably. Four timers are exploited for motor driving, camera driving, IR signal capturing and speed capturing. The interrupt mechanism is used for real-time capturing. In order to debug the software system, an UART based command line interface is designed. By using the PC command line, each parameter can be on-line adjusted and

---

[1] All the hardware schematics and software codes will be available for download after the paper is accepted.

stored. Moreover, with the assistance of script programs, cameras and other related devices can be controlled automatically. The motor close loop control method is based on PID algorithm [12], which can control the speed and distance of the movement.

The current version of software utilizes only half of the CPU resources, which means the software is able to be updated in the future. More functions can be implemented on this platform. The source code of the proposed platform is available for download at http://www.mmtlab.com/platform.

## 3 Test and Related Experiments

### 3.1 Tests

After PCB customization and system assembly, the proposed platform was tested. The PID parameters were adjusted to make the platform move smoothly and accurately. The movement of this platform can satisfy the requirement of variety of experiments. It is worth mentioning that the software could be modified based on specific requirement of different experiments. Therefore, this platform is flexible and programmable.
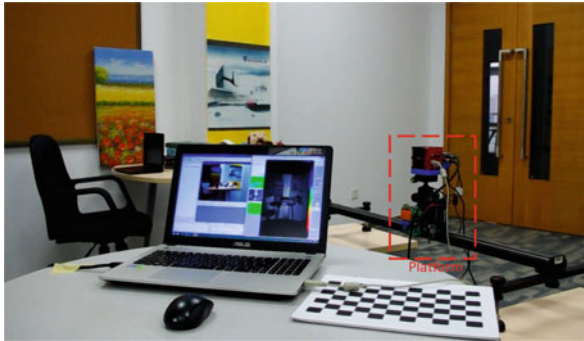


**Fig. 8.** A related experiment based on the proposed platform

### 3.2 Related Experiments and Results

This platform has been used to conduct two experiments related to some proposed methods [10]. Moreover, some new methods are being developed based on this platform. Figure 8 shows a related experiment using this platform. In this case, we use one texture camera and one depth camera, while the platform is moving. The video sequences of texture and depth are captured synchronously with the motion information. By using the produced sequence, Fig. 9 reports that a proposed method which exploits the camera global motion information

is better than the baseline codec (i.e., JM software [13] without using motion information). The BD-Rate is -11.87 %, while the BD-PSNR is 0.49 dB.

The other experiment is depth map coding using motion sensor information. This experiment verified that the proposed global motion information assisted depth map coding is more efficient than the baseline codec (i.e., JM software). The BD-Rate is -41.12 %, and the BD-PSNR is 2.04 dB. The experiments based on the platform indicate that the platform can be beneficial to various emerging video coding scenarios (Fig. 10).
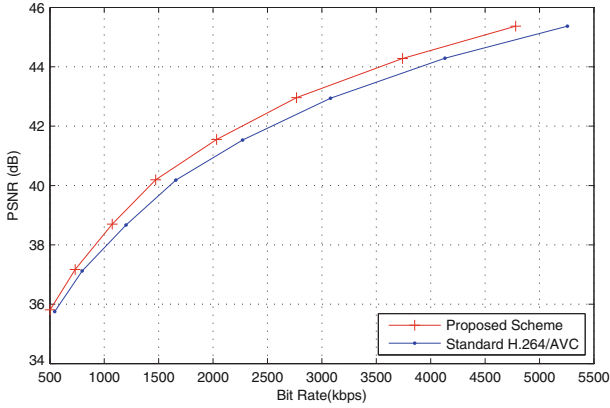


**Fig. 9.** The RD-performance of one experiment of texture coding based on camera global motion and depth information
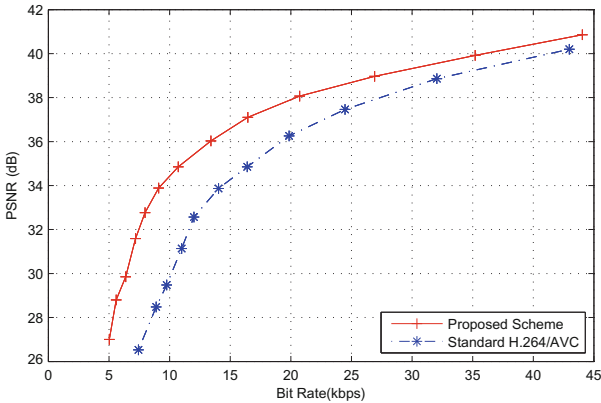


**Fig. 10.** The RD-performance of one experiment of depth map coding based on camera global motion information

# 4   Conclusions

This paper introduced a flexible and programmable camera control and video data acquisition platform. This platform could control the movement of cameras and record the motion information. Multiple types of cameras could be driven and captured synchronously. The sampling of power grid is exploited to adjust the frame rate of video capturing in order to reduce the flicker effect. It is worth mentioning that the hardware schematic, layout and the software source codes are published and available for download.

With the assistance of this platform, we completed some video coding experiments and obtained positive results. In the future, this platform will be updated continuously. More functions will be implemented on it, while the performance of the existing functions will be improved.

# References

1. CISCO White Paper, Cisco Visual Networking Index: Forecast and Methodology, 20132018. http://www.cisco.com/
2. Angelino, C.V., Cicala, L., Persechino, G., Baccaglini, E., Gavelli, M., Raimondo, N.: A sensor aided H.264 encoder tested on aerial imagery for SFM. In: IEEE International Conference on Image Processing (ICIP), Paris, France, 10/2014 (2014)
3. Chen, X., Zhao, Z., Rahmati, A., Wang, Y., Zhong, L.: Sensor-assisted video encoding for mobile devices in real-world environments. IEEE Trans. Circuits Syst. Video Technol. **21**(3), 335–349 (2011)
4. Merkle, P., Smolic, A., Muller, K., Wiegand, T.: Multi-view video plus depth representation and coding. In: IEEE International Conference on Image Processing, ICIP 2007, vol. 1, pp. I-201. IEEE (2007)
5. Fehn, C.: Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. In: International Society for Opticsand Photonics Electronic Imaging 2004, pp. 93–104 (2004)
6. Hannuksela, M., Chen, Y., Suzuki, T., Ohm, J.-R., Sullivan G., (ed.) AVC draft text 8. In: JCT-3V document JCT3V-F1002, vol. 16 (2013)
7. Chen, Y., Hannuksela, M.M., Suzuki, T., Hattori, S.: Overview of the MVC+ D 3D video coding standard. J. Vis. Commun. Image Represent. **25**(4), 679–688 (2014)
8. Bucci, G., Fiorucci, E., Landi, C.: A digital instrument for light flicker effect evaluation. IEEE Trans. Instrum. Meas. **57**(1), 76–84 (2008)
9. Qiao, S., Zhang, Y., Wang, H.: PI-frames for flickering reduction in H.264/AVC video coding. In: 2012 International Conference on Computer Science Service System (CSSS), pp. 1551–1554, August 2012
10. Jin, Z., Tillo, T., Cheng, F.: Depth-map driven planar surfaces detection. In: IEEE Visual Communications and Image Processing Conference (IEEE VCIP), Valletta, Malta, 12/2014 (2014)
11. MESA IMAGING. SR4000. http://www.mesa-imaging.ch/products/sr4000/
12. Shumway-Cook, A., Woollacott, M.H.: Motor Control: Theory and Practical Applications, vol. 157. Williams & Wilkins Baltimore, Baltimore (1995)
13. HHI Fraunhofer Institute, H.264/AVC Reference Software. http://iphome.hhi.de/suehring/tml/download/