3D shape measurement accelerated by GPU

Zhao Yalong, Liu Shouqi, Zhang Qican

(School of Electronics and Information Engineering, Sichuan University, Chengdu 610064, China)

Abstract: Driven by the increasing demands of the general purpose in computation and image display, Graphics Processing Unit (GPU) has been developed and used in many fields, such as medical field, scientific calculations, image processing etc.. But, its application in 3D shape measurement is still a beginning. In this paper, two 3D shape measurement systems based on Fourier Transform Profilometry (FTP) and tri-frequency heterodyne method were implemented with Compute Unified Device Architecture (CUDA) technology to speed up their 3D shape construction of a measured static or dynamic object. In the first 3D shape measuring system based on tri-frequency heterodyne method, a high-speed digital projection module and a synchronously triggered camera were used to record 12 deformed fringe images on the surface of a small object. The experimental result demonstrates that the efficiency of the unwrapping phase calculation by GPU is improved 2 089 times than that of CPU for doing same task on 12 images with 1 360 pixel×1 024 pixel each. In the second system based on FTP, only one deformed fringe image was recorded by a camera, then transferred into GPU and processed by the programmed CUDA algorithm to restore the corresponding 3D shape. Compared with the traditional processing method by CPU, the time consumption of FTP method completed by GPU is shortened 27 times for a 1 024 pixel×1 280 pixel image.

Key words: 3D shape measurement; GPU; fringe projection; Fourier transform profilometry; tri-frequency heterodyne

CLC number: TN247 **Document code:** A **DOI:** 10.3788/IRLA201847.0317003

GPU 加速三维面形测量

赵亚龙,刘守起,张启灿

(四川大学 电子信息学院,四川 成都 610064)

摘 要: 随着通用计算和图形显示需求的不断增加,图形处理器(Graphics Processing Unit,GPU)在医学、科学计算、图像处理等领域得到了广泛的应用。但它在三维测量领域的应用还只是一个开始。文中基于傅里叶变换轮廓术(Fourier Transform Profilometry,FTP)和三频外差法设计了两套三维测量系统,并利用计算统一设备架构(Compute Unified Device Architecture,CUDA)方法,加速了静态或动态物体的三维重建。在三频外差测量系统中,需要利用高速数字投影模块和相机,同步触发采集小视场表面的 12 个变形条纹图,然后对图像数据进行处理。实验结果表明:对 12 幅 1 360 pixel×1 024 pixel 大小的图像进行相位展开运算,GPU 方法比 CPU 方法的效率提高了 2 089 倍。在基于 FTP 方法的测量

收稿日期:2017-10-05; **修订日期**:2017-11-15

基金项目:国家自然科学基金(61675141);国家重大科学仪器设备开发专项(2013YQ490879)

作者简介:赵亚龙(1990-),男,硕士生,主要从事三维面形测量方面的研究。Email:616475529@qq.com

导师简介:张启灿(1974-),男,教授,博士,主要从事三维传感、动态三维测量等方面的研究。Email:zqc@scu.edu.cn

系统中,摄像机只需记录一幅变形条纹图,然后拷贝到显存中,并用 CUDA 编程的算法进行处理,进而重建出物体的三维面形。基于 GPU 的 FTP 方法对一幅 1 024 pixel×1 280 pixel 大小的图像进行计算,其计算时间比 CPU 方法缩短了 27 倍。

关键词:三维面形测量; GPU; 条纹投影; 傅里叶变换轮廓术; 三频外差

0 Introduction

Compared with the 2D image information, the 3D shape distribution can more completely and truly reflect the objective object and provide a greater amount of information. Therefore, 3D shape measurement becomes a search hotspot. Optical 3D measurement, with many advantages such as non-contact, fast measurement speed, high precision, simple device, easy to operate, etc., is one of the most important methods of 3D shape measurement. It is widely used in face recognition, industrial inspection and other fields[1-3]. Usually, an active 3D shape measuring system uses a well designed structured light during the measuring process. The employed structured light is projected onto the surface of the object, and will be modulated by the 3D information of the tested object. The deformable structured image pattern can be acquired through an imaging device from another direction. Then through the demodulation processing of Fourier fringe analysis or phase shifting analysis, the 3D shape of the object can be reconstructed. This kind of active 3D sensing method has high accuracy and sensitivity, so it is commonly used in actual measurement.

Among the existing phase demodulation methods, Fourier fringe analysis method needs only one single frame image to recover its corresponding wrapped phase distribution by Fourier transformation and filtering operation, and then a point by point phase unwrapping algorithm will be performed on 2D space to obtain its continuous phase, which will be used to thereby reconstruct the tested 3D shape. The phase shifting analysis method is another popular one, multi-frame phase shift fringes with different frequency will be used to obtain their corresponding wrapped phase distribution. In accordance with their intrinsic relationship, the 3D shape will be reconstructed pixel by pixel along the time axis. If these two kinds of algorithms are implemented in CPU, their time-consumption will be large which can′t meet the needs of fast and even high speed 3D shape measurement.

With the development of artificial intelligence, virtual reality, face recognition and industrial detection, the requirement for the accuracy and speed of 3D data acquisition is increasing. The required resources and computing time is also increasing, which brings great challenges to the research in these fields. Unified computing device architecture published by NVIDIA can efficiently make use of GPU strong processing power and huge memory bandwidth in calculation other than graphics rendering, it has been widely used in many fields of modern science. But in the field of 3D shape measurement, it′s only in the beginning. Song Zhang et al. used two-plus-one phase-shifting algorithm based on GPU to process the 266 K image which achieved the speed of 30 fps in 2006[4]. Nikolaus Karpinsky et al. achieved the real-time 3D topography measurements on portable devices (laptops) in 2014[5]. Lu Jin et al also applied the GPU to the surface morphology measurement system[6].

Based on the analysis of two commonly used methods of 3D measurement algorithm, FTP and tri-frequency heterodyne method, and combined with the characteristics of the GPU programming model, the original CPU algorithm was parallel designed and relevant experiments were carried out as well. The experimental results show that the parallel algorithm based on GPU has greatly reduced the time-consumption than the original serial algorithm executed in CPU.

1 Principle

1.1 CUDA programming model

Typically, CPU can only compute one task at one time, while GPU can handle multiple tasks simultaneously, so GPU can accelerate the data calculation speed[7-9]. Figure 1 shows the comparison of parallel computation and serial computation. In the CUDA programming model, CPU acts as a Host and GPU as a co-processor or device, and GPU is mainly responsible for the implementation of highly threaded parallel processing tasks, while CPU is responsible for processing a series of logically controlled things and serial computing. When designing a CUDA program, the parallel and non-parallel parts of the program must be defined first, the parallel computing parts of the CUDA framework is in the kernel function, and the control process is on the CPU side. A complete CUDA program includes a series of parallel functions of the device kernel function and the serial computing steps of the host.

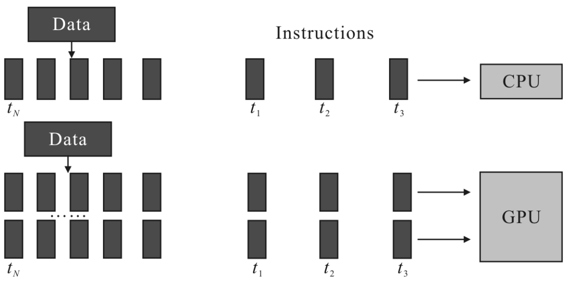

Fig.1 Comparison of the serial computation in CPU and the parallel computation in GPU

A simple CUDA programming model consists of the following steps:

(1) Initialize the data on the CPU; (2) Transmit the data to GPU; (3) Run the kernel function to process the data; (4) Return the results to the CPU.

1.2 Principle of phase calculation based on tri-frequency heterodyne method

Tri-frequency heterodyne method is one of the mainstream 3D measurement methods in the temporal phase unwrapping[10]. This method has the advantages of fewer projected fringes and high measurement precision. Because most of the tri-frequency heterodyne calculations are point-to-point parallel calculations, this method is particularly suit for GPU-accelerated processing.

The tri-frequency heterodyne method need three wrapped phase $\phi_k(x,y)(k=1,2,3)$ with different frequencies, which can be calculated by four-step phase shifting method, as shown in Eq.(1):

$$\phi_k(x,y)=\arctan\left(\frac{I_4(x,y,k)-I_2(x,y,k)}{I_1(x,y,k)-I_3(x,y,k)}\right) \tag{1}$$

Where I_n $(x,y,k)(n=1,2,3,4)$ represent intensity values of each image.

In the heterodyne method a lower frequency phase function $\Phi_b(x,y)$ is superposed by two different phase functions $\phi_1(x,y)$ and $\phi_2(x,y)$, as shown in Eq.(2), where λ_1, λ_2 and λ_b are the corresponding frequencies of the phase functions ϕ_1 (x,y), ϕ_2 (x,y) and $\Phi_b(x,y)$, respectively.

$$\lambda_b=\frac{\lambda_1\lambda_2}{\lambda_1-\lambda_2} \tag{2}$$

The heterodyne principle can be used to solve the problem of phase unwrapping[11]. The frequencies of the phase functions $\phi_1(x,y)$, $\phi_2(x,y)$ have to be chosen in a way that the resulting phase function $\Phi_b(x,y)$ is unambiguous over the field of view, that is to say $\lambda_b=1$. Due to the construction of the phase function $\Phi_b(x,y)$, the ratio R_1 is always constant. The unwrapped phase of $\phi_1(x,y$, $\Phi_1(x,y)$ can be resulted by Eq.(3):

$$\Phi_1(x,y)=\phi_1(x,y)+O_1(x,y)\times2\pi \tag{3}$$

In Eq.(3), $O_1(x,y)=\mathrm{INT}\left(\frac{\Phi_b(x,y)\times R_1-\phi_1(x,y)}{2\pi}\right)$, INT is the inter operation.

In this method the wrapped phase is obtained by Eq.(1), which is a subtraction and division operation for each pixel respectively. In the process of the unwrapped phase calculation, the operation of the heterodyne and phase unwrapping is performed on each pixel along its own timeline. Through the analysis of the above steps, it can be seen that the computation of pixels is independent of each other.

Therefore, compared with the serial computing of CPU method, using parallel computation of GPU can greatly reduce the time consumption. The flow chart Fig.2 shows the whole data process.

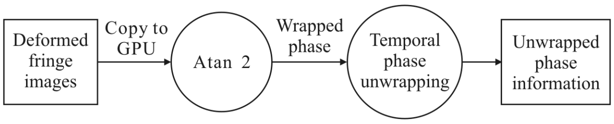

Fig.2 Phase unwrapping process in tri-frequency heterodyne method

2.3 Principle of phase calculation based on Fourier fringe analysis

In 1983, M. Takeda and K. Mutoh first proposed the FTP for the auto measurement of 3D object shapes based on grating projection[12-13]. In the FTP, a standard sinusoidal fringe is projected to the surface of the measured object, and the deformed fringe image captured by a camera which can be expressed as Eq.(4):

$$I(x,y)=a(x,y)+b(x,y)\cos[2\pi f_0 x+\phi(x,y)] \tag{4}$$

Where $I(x,y)$, $a(x,y)$, $b(x,y)$, $\phi(x,y)$ represent intensity values of each pixel, background illumination, object surface reflectance, and phase distribution respectively. And f_0 is the frequency of the projected fringe pattern.

After Fourier transform, the Fourier spectrum of $I(x,y)$, expressed as $G(f_x,f_y)$ in Eq.(5), can be obtained.

$$G(f_x,f_y)=A(f_x,f_y)+Q(f_x-f_0,f_y)+Q^*(f_x-f_0,f_y) \tag{5}$$

In Eq.(5), $A(f_x,f_y)$ represents zero component of the spectrum, $Q(f_x-f_0,f_y)$ and $Q^*(f_x-f_0,f_y)$ represent the fundamental components. With a proper digital filter, one of the fundamental components can be extracted. Then the wrapped phase can be obtained by inverse Fourier transform. And after phase unwrapping, the unwrapped phase $\phi(x,y)$ can be obtained.

The same gating image is projected to a reference plane and carried out the same operations on its image, both the unwrapped phase of the reference plane $\phi_0(x,y)$ and the relative phase difference can be obtained and expressed in Eq.(6).

$$\Delta\phi(x,y)=\phi(x,y)-\phi_0(x,y) \tag{6}$$

The mapping relation between the phase difference $\Delta\phi(x,y)$ and the height $H(x,y)$ expressed as Eq.(7) can be finally established through system calibration.

$$H(x,y)=-\frac{L\Delta\phi(x,y)}{2\pi f_0 d} \tag{7}$$

In Eq.(7), L is the distance between the entrance pupil of the imaging system and the reference plane, and d is the distance between the entrance pupil of the imaging system and the exit pupil of the projector.

Fourier transform is a typical divide and conquer algorithm, which divides a complex problem into two or more identical or similar sub-problems, and then sub-problems into smaller sub-problems until the final sub-problems can be simple solved directly. Therefore, the Fourier transform can be parallel calculated in GPU. In this method, the phase unwrapping was done by CPU because the spatial phase unwrapping is an integral accumulation process of point-by-point scanning. The flow chart Fig.3 shows the whole data process.

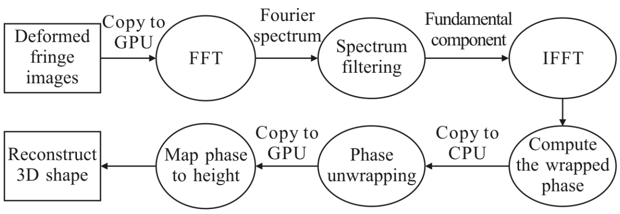

Fig.3 FTP data process accelerated by GPU

3 Experimental results

3.1 First experiment on phase calculation based on tri-frequency heterodyne method

The experiment system was built based on the principle of triangulation, the main hardware devices were a digital projector module and IDS industrial camera (The resolution is 1 360 pixel×1 024 pixel). The experiment setup and a measured coin are shown in Fig.4.

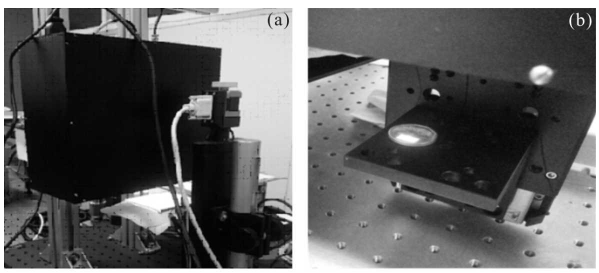

Fig.4 (a) Experiment setup and (b) a measured coin

The digital projector projected 12 fringe images of 3 frequencies (1/42, 1/49, 1/57) to the tested coin′s surface. Fig.5(a) shows the 4 deformed fringe images of λ=1/42. All the deformed fringe images synchronously recorded by the CCD camera, and then the corresponding 3-D shape of the tested coin can be reconstructed and shown in Fig.5(b).

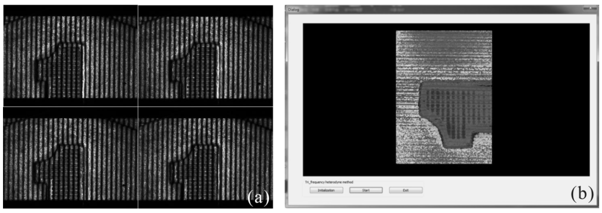

Fig.5 (a) Four deformed fringe images of λ=1/42, (b) reconstructed result of the coin

The data analysis process was carried out under the hardware condition of NVIDA GTX580 and Inter Core i7. For comparing the reconstruction result, the same fringe patterns were processed in CPU and GPU respectively with the same algorithm. Their time consumptions were counted using the Query Performance Counter method as shown in Tab.1. It can be seen that the speedup ratio (CPU time-consumption/GPU time-consumption) of the process (including the phase calculation and unwrapping) by GPU is 2 089 for the same task on 12 images with 1 360 pixel×1 024 pixel each.

Tab.1 Comparison of the time consumptions of tri-frequency heterodyne method by CPU and GPU

	Time-consumption on CPU/ms	Time-consumption on GPU/ms	Speedup ratio
Wrapped phase calculation	87.062	0.071	1 226
Temporal phase unwrapping	188.688	0.061	3 093
Total time	275.750	0.132	2 089

3.2 Second experiment based on FTP

In order to verify the feasibility of the GPU in FTP method and the acceleration of data processing, the second experiment took a static semi-sphere shaped object as the measured object. The main hardware device was a high-speed CMOS camera FASTCAM Mini UX100 and a digital projector module DLP Lightcrafter 4500. The experiment setup is shown in Fig.6. The digital projector has more than one million micromirrors with a projection rate of 4 225 Hz at 1 bit and 120 Hz for 8 bit. The camera worked at 1 000 fps with the resolution of 1 024 pixel× 1 280 pixel.

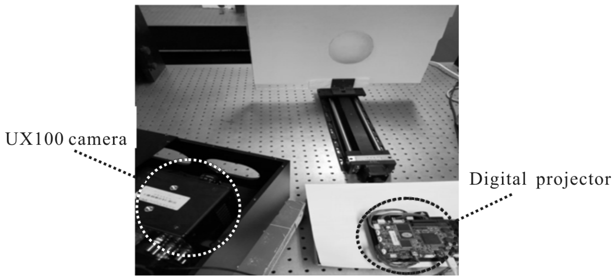

Fig.6 Experiment setup based on FTP

The deformed fringe image (shown in Fig.7(a)) recorded by the CMOS camera was transferred to the GPU, and then through Fourier transform, filter and inverse Fourier transform, the wrapped phase was obtained. The spatial phase unwrapping is a typical integrating process which is difficult to be deal with by parallel computation, so it was done by CPU in this experiment. After the phase being unwrapped in CPU, the measured object′s 3D shape information was restored with the loaded calibration parameters and the mapping relation. Figure 7(b) shows the reconstructed result.

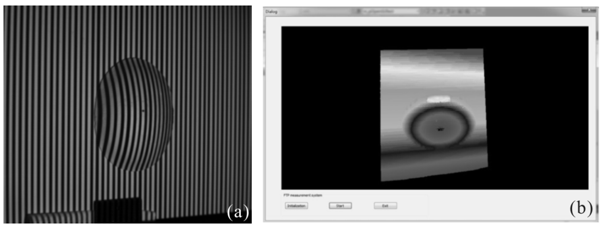

Fig.7 (a) Deformed fringe image, (b) reconstructed result of the semi-sphere shaped object

The hardware condition of the computer was the same as the first experiment. It also used Query

Performance Counter as a counting tool to record the processing time in the GPU and CPU for several major steps of the FTP method, and the result is shown in Tab.2. For a 1 024 pixel×1 280 pixel image, the total time consumption by CPU is 176.484 ms while 29.949 ms by GPU, the speedup ratio is 27. In this experiment, the process of phase unwrapping was executed in CPU.

Tab.2 Comparison of the time consumptions of FTP method by CPU and GPU

	Time-consumption by CPU/ms	Time-consumption by GPU/ms	Speedup ratio
FFT	21.888	1.148	19
Spectrum filtering	6.317	0.018	351
IFFT	31.649	0.051	621
Wrapped phase calculation	16.477	0.005	3 295
Phase unwrapping (Completed in CPU only)	24.433	–	–
Copy unwrapped phase data to GPU	–	4.260	–
Map phase to height	75.720	0.034	2 227
Total time (Without phase unwrapping)	176.484 (152.051)	29.949(5.516)	27

On the basis of this experimental platform, a dynamic 3D shape measurement for a vibrating loudspeaker was completed too. The high speed CMOS camera was used to record the deformation fringes at different times and then the software system processed them. The reconstructed result can demonstrate the loudspeaker′s 3D changing shape at different time, which is not shown in this paper due to the manuscript length limitation.

4 Conclusions

In this paper, in view of the demand of measuring speed in 3D measurement field, the application of the GPU parallel computing in 3D measurement was analyzed in detail and two sets of experiments were designed. The experimental result shows the feasibility of CUDA technology in FTP and tri-frequency heterodyne measurement methods, and the speed of the data processing is obviously improved. Compared with the traditional processing method by CPU, the time consumption of FTP method is shortened 27 times for a 1 024 pixel×1 280 pixel image and the unwrapped phase calculation in tri-frequency heterodyne method by GPU is improved 2 089 times than that of CPU for doing same task on 12 images with 1 360 pixel×1 024 pixel each. Both two methods above mentioned have been made some theoretical support for parallel computing in 3D shape measurement filed, which make a real-time 3D shape measurement system possible. In the future work, a feasible 3D measurement algorithm running in GPU is urging to be proposed to develop a real-time measurement system for dynamic object.

For the phase unwrapping in two phase calculation methods in this paper, the tri-frequency heterodyne is a temporal-based phase unwrapping method, which can be parallel calculated by GPU. However, the second method is based on Fourier analysis, so its phase unwrapping is a typical integrating process and not suitable to be deal with the parallel computation by GPU. This problem can be bypassed by setting up a lookup table of the relationship between the wrapped phase and the restored height distribution.

References:

[1] Tian Qingguo, Ge Baozhen, Du Pu, et al. Measurement of human figure size based on laser 3D scanning [J]. ***Optics and Precision Engineering***, 2007, 15(1): 84-88. (in Chinese)

[2] Tang Wei, Ye Dong. 3D computer vision measurement systems [J]. ***Infrared and Laser Engineering***, 2008, 37(S1): 328-332. (in Chinese)

[3] Ye Haijia, Chen Gang, Xing Yuan. Stereo matching in 3D

measurement system using double CCD structured light [J]. ***Optics and Precision Engineering***, 2004, 12(1): 71–75. (in Chinese)

[4] Zhang S, Royer D, Yau S T. GPU–assisted high-resolution, real-time 3–D shape measurement [J]. ***Optics Express***, 2006, 14(20): 9120–9129.

[5] Nikolaus Karpinsky, Morgan Hoke, Vincent Chen, et al. High-resolution, real-time three-dimensional shape measurement on graphics processing unit [J]. ***Optical Engineering***, 2014, 53(2): 024105–1–8.

[6] Lu Jin. The surface morphology measurement system based on GPU [D]. Hangzhou: Zhejiang University, 2011. (in Chinese)

[7] Zhang Shu, Chu Yanli, Zhao Kaiyong, et al. GPU Performance Computing of CUDA [M]. Beijing: China Waterpower Press, 2009: 11–13.

[8] Chen Qian, Qiu Yuehong, Yi Hongwei. Star image registration algorithm based on GPU parallel program design [J]. ***Infrared and Laser Engineering***, 2014, 43(11): 3756–3761. (in Chinese)

[9] Wang Xinhua, Wang Xiaokun. Real time image mosaic of the transient gigapixel imaging system [J]. ***Chinese Optics***, 2015, 8(5): 785–793. (in Chinese)

[10] Huntley J M, Salder H O. Shape measurement by temporal phase unwrapping: comparison of unwrapping algorithms [J]. ***Measurement Science Technology***, 1997, 8(9): 986–992.

[11] Carsten Reich, Reinhold Ritter, Jan Thesing. White light heterodyne principle for 3D–measurement [C]//SPIE, 1997, 3100: 236–344.

[12] Takeda M, Mutoh K. Fourier transform profilometry for the auto measurement of 3–D object shapes [J]. ***Applied Optics***, 1983, 22(24): 3977–3982.

[13] Lei Gundong, Lu Yinhuan, Wang Ruli. Adaptive main frequency bandpass filters used in Fourier transform profilometry [J]. ***Chinese Journal of Optics and Applied Optics***, 2010, 3(3): 245–251. (in Chinese)