Received: 29 December 2024 | Revised: 19 January 2025 | Accepted: 2 February 2025 | Published online: 6 February 2025

International Journal of Advance in Applied Science Research, Volume 4, Issue 2, 2025
https://h-tsp.com/

Research on the Application of GPU Parallel Computing in Image Processing

Xiaoyu Liu*

School of Computer Science, Xianyang Normal University, Xianyang, Shaanxi 712000, China
**Author to whom correspondence should be addressed.*

Abstract: ***The development of the Internet has revolutionized information technology, propelling it into an era of unprecedented advancement. One significant area that has witnessed a transformative evolution is the processing of video images. With the advent of new technologies, the methods for manipulating and enhancing these images have become increasingly diversified, catering to a wide array of applications and industries. Each processing technique boasts its unique set of characteristics, tailored to meet specific needs and deliver optimal results. In this rapidly evolving landscape, it is crucial to stay abreast of the latest innovations and advancements. With this in mind, this article embarks on a detailed exploration of one such cutting-edge image processing method: the C++ language image computing method. This method stands out for its efficiency and convenience, offering a robust platform for handling intricate image processing tasks with ease. The C++ language image computing method shines particularly bright when it comes to showcasing the benefits of GPU parallel addition. By leveraging the powerful parallel processing capabilities of GPUs, this method significantly enhances computational speed and accuracy. To illustrate this point, the article delves into the parallelization of the Gaussian blur algorithm and color negative film processing. By analyzing these specific applications, it becomes evident how GPU parallel addition can drastically reduce processing time while maintaining the integrity and quality of the results. The implications of adopting such an advanced image processing method are far-reaching. It not only elevates the efficiency of image manipulation tasks but also paves the way for innovative applications in fields such as photography, film production, medical imaging, and beyond. Therefore, it is imperative that we vigorously promote and apply this C++ language image computing method. By embracing this technology, we can unlock new possibilities, drive progress, and ultimately, revolutionize the way we perceive and interact with visual information in the digital age.***

Keywords: GPU parallel computing; Image processing; Applied research; Chip; Kernel.

1. Introduction

In recent years, with the development of information technology, the development of graphics processing chip technology has also become increasingly mature.

At present, GPUs have become relatively powerful computing devices, especially in programmable image processing, which has surpassed traditional processing methods by tens of times or even more, and has also made programmable image processors the main force in the development of computer technology. And it broke the deadlock of CPU dominance in the past, completely changing the direction and scale of the development of the computer industry.

Compared to CPUs, GPUs have a stronger market power and therefore far exceed the development speed of CPUs. In terms of development speed, GPU development is far ahead, surpassing Moore's Law of computing and becoming an important node in the development of computer technology in the new era. Although parallel computing started relatively late, it has still received some attention, especially with the increasing maturity of multi-core development environments, and the development speed of parallel computing has also made some progress. This article provides a detailed analysis of the application of GPU parallel computing in image processing, in order to better improve computing power and promote the development and maturity of computer technology.

Recent advancements in various fields, including urban building energy optimization, 3D point cloud object recognition, crystal system classification, e-commerce chatbots, supply chain and digitization impacts on environmental technologies, clinical trial strategies for anti-HER2 drugs, subgraph matching, subgraph enumeration, in-network content adaptation, 5G RAN slicing, domain generalization in 3D human pose estimation, manufacturing firms' green innovation and organizational performance, IoT-based 3D pose estimation for athletes, and online update methods for retrieval-augmented generation (RAG) models.

Zheng et al. [1] proposed a Triz method for urban building energy optimization, utilizing a GWO-SARIMA-LSTM forecasting model. The study aimed to improve energy efficiency in urban buildings. Lyu et al. [2] optimized CNNs for rapid 3D point cloud object recognition, enhancing the accuracy and speed of recognition processes. Yin et al. [3] applied deep learning techniques to classify crystal systems in lithium-ion batteries, contributing to advancements in battery technology. Luo et al. [4] enhanced e-commerce chatbots using Falcon-7B and 16-bit full quantization, improving their performance and efficiency. Liu et al. [5] investigated the impact of supply chain and digitization on the development of environmental technologies in G7 nations, revealing the significance of inflation and consumption.

Li [6] focused on optimizing clinical trial strategies for anti-HER2 drugs using Bayesian optimization and deep learning, aiming to improve drug development processes. Jin et al. [7] presented FAST, an FPGA-based subgraph matching system for massive graphs, demonstrating its efficiency and scalability. Yang et al. [8] introduced HUGE, an efficient and scalable subgraph enumeration system, further advancing subgraph-related research. Chen et al. [9] proposed Octopus, an in-network content adaptation system to control congestion on 5G links, enhancing network performance. In another study, Chen et al. [10] presented a channel-aware 5G RAN slicing system with customizable schedulers, addressing the challenges of network slicing.

Peng et al. [11] developed a dual-augmentor framework for domain generalization in 3D human pose estimation, improving the robustness and accuracy of pose estimation algorithms. Yan et al. [12] examined the relationship between CEO power and manufacturing firms' green innovation and organizational performance, revealing important insights for business management. Ren et al. [13] applied IoT-based 3D pose estimation and motion optimization for athletes, using C3D and OpenPose, providing valuable tools for sports training and performance analysis. Fan et al. [14] researched the online update method for retrieval-augmented generation (RAG) models with incremental learning, contributing to advancements in natural language processing.

2. Parallel Computing Method

Parallel computing, as a new discipline, although developed and started relatively late, has received a good response after being put into use, and therefore deserves to be vigorously promoted and applied. Compared to serial computing, parallel computing is mainly divided into two types: temporal parallel

and spatial parallel. Among them, time parallelism, also known as pipeline technology, and spatial parallelism are used to perform calculations.

In response to the needs of this topic, we mainly study spatial parallelism in parallel computing. Space parallelism can be divided into data parallelism and task parallelism from the perspective of program and algorithm designers. Data parallelism is the process of dividing an entire task into several small tasks and completing them one by one, in order to improve processing efficiency. In data parallel algorithms, there are mainly four types of algorithms:

2.1 By dividing multi-core CPUs into 4 or 8 cores, the more cores there are, the stronger the computing power.

2.1.1 Main advantages of multi-core CPU:

(1) The hardware price is relatively low.
(2) Software development is quite challenging.

(3) The development content is too complex, but the resource content is relatively rich.

(4) The maintenance cost of multi-core CPUs is relatively low, and software can run on higher generation CPUs without any modification, greatly saving maintenance costs.

2.1.2 Main drawbacks of multi-core CPUs:

(1) Due to certain limitations in computing power, it cannot meet the requirements for large-scale and massive data processing.

(2) Without any expansion capability, it must be updated and replaced in a timely manner.

2.2 Multiple PCs simultaneously computing a task

Multiple PCs, also known as cloud computing, are currently one of the most commonly used computing methods.

2.2.1 The computing advantages of cloud computing lie in:

(1) The security factor of data storage centers is relatively high, and the possibility of data loss or virus invasion is small.

(2) The device requirements for users are relatively low, and for users, it is more convenient to use.

(3) Ability to achieve resource and data sharing among different users.

2.2.2 The computing drawbacks of cloud computing are:

(1) The computational requirements are directly proportional to the hardware cost, which means that the higher the hardware configuration, the stronger the processing power.

(2) The content in software development is quite complex, which is not conducive to development.

(3) Due to the fact that parallel computing belongs to high-end applications and software resources are relatively scarce, the development resources for this computing method are also relatively scarce.

(4) The maintenance cost is too high, and professional maintenance personnel are needed.

2.3 Multi CPU mainframe

2.3.1 Advantages of Multi CPU Mainframe:

The advantage lies in its powerful performance, which can meet all needs.

2.3.2 Disadvantages of Multi CPU Mainframe:

(1) The manufacturing process has a relatively large volume and high manufacturing costs.

(2) Software development is difficult and requires a high level of professional expertise from professionals, who need to have some development experience.

(3) Lack of software resources.

(4) The subsequent maintenance and operation costs are too high, and some problems cannot be resolved in a timely manner.

2.4 In terms of collaborative processing

2.4.1 Advantages:

In terms of collaborative processing, CPUs are more inclined towards logical computing, while GPUs are more inclined towards mathematical computing. Therefore, these two processing methods can only have strong computing power and greatly improve computing efficiency on the basis of parallel computing. At the same time, there are low hardware costs, convenient hardware upgrades, low maintenance costs, and low power consumption.

2.4.2 Disadvantages:

(1) For developers, it is necessary to learn new programming languages to cope with the difficulties encountered in software development.

(2) As a new discipline, parallel computing has certain development space in both software resources and professional level.

(3) Collaborative processing still lacks control capabilities and belongs to non universal design.

(4) There are clear procedures and requirements for the calculation method.

(5) The calculation process consumes a relatively long time and has certain advantages in processing a large amount of data, but the efficiency in processing some small data is not significant.

3. Application Analysis of Parallel Algorithms in Image Processing

3.1 Parallelization of Gaussian Fuzzy Algorithm

Gaussian blur algorithm is one of the commonly used calculation methods in image processing. Compared to other calculation methods, Gaussian blur algorithm has the advantages of reducing image noise and lowering detail levels. But it also has certain limitations, that is, this algorithm cannot be effectively transformed in parallel algorithms. That is to say, the Gaussian fuzzy algorithm must strictly follow the reconstruction program processed by the CUDA multi-threaded framework. Especially in the process of blurring images, it is necessary to always ensure that there is no correlation between Gaussian matrices, in order to achieve the goal of separate processing. Different tasks are divided into several small modules and then calculated to improve computational efficiency.
In the calculation process, the Gaussian blur algorithm first imports the original graphics of the image to be processed, and then divides the image pixels into horizontal and vertical blocks for processing. Finally, the processing results of all segmented small task parts are integrated to complete the Gaussian fuzzy algorithm and improve its efficiency.

3.2 Parallelization of Color Negative Film Processing Algorithms

The color negative film algorithm requires a lot of time from Xiao Han in both processing methods and processes, mainly due to the color negative film algorithmIn parallel computing, it is necessary to process each pixel first before conducting analysis. However, in terms of operation, the processing method of color negative film method is relatively simple, that is, after reading in the image, the image is spread throughout the pixels to obtain corresponding data information, and then these data information are integrated through color negative film method to obtain the final value, which can generate a new image.

Compared to the Gaussian blur algorithm, the color negative film method is independent in terms of calculation and image processing, that is, during the processing, the pixel value of a certain point changes, while other numerical points do not change. Therefore, the color negative film method is more suitable for GPU multi-threaded parallel computing, integrating image data information into new images.

3.3 Parallelization of Transparent Merge Processing Algorithm

The transparent merging algorithm aims to achieve a semi transparent effect on the image. The parallelization formula for the transparent merge processing algorithm is shown in formula (1). We set the opaque image to be processed as A, the transparent object as B, and after applying the transparency merging algorithm, we set the resulting new color image as C. After calculation, B will obtain a numerical value, which should be controlled within the range of 0-1. Among these values, 0 indicates that the object is completely transparent, and 1 indicates that the object is completely opaque.

$$\begin{cases} R(C) = alpha \times R(B) \times (1 - alpha) \times R(A) \\ G(C) = alpha \times G(B) \times (1 - alpha) \times G(A) \\ B(C) = alpha \times B(B) \times (1 - alpha) \times B(A) \end{cases} \quad (1)$$

From the formula, we can see the components of pixel color represented by different values. And the transparent merging algorithm is relatively simple in terms of hybrid processing, which can obtain semi transparent effects such as smoke, firelight, and shadows.

At the same time, in the GUDA framework, in order to simultaneously process two images that need to be mixed and transparent, one is to first determine the storage location of the images; The second is to select the range of images that need to be processed; The third is to divide the parts of the image that need to be processed into cells and control these cells through the CPU; The fourth is to process and operate these already allocated cells separately through a multi-threaded processor.

Compared to other processing methods, this processing method is relatively simple in operation and does not require excessive logical control during the operation process. Even when facing different processing objects, the overall processing methods are similar. At the same time, different processing objects are not related to each other, so the advantage of transparent merging algorithm is that it has great convenience and efficiency.

4. CONCLUSION

In summary, the rapid development of information technology has promoted the development of computer technology, which has been widely applied in both daily life and image processing. At the same time, in this context, people's requirements for the quality and efficiency of image processing are constantly increasing. This article provides a detailed analysis of commonly used image processing methods, such as Gaussian blur processing, color negative film processing, and transparent merging processing. These image processing methods all need to be based on the GUDA programming framework, and then use GPU multi-threaded module parallel processing methods to efficiently process images, which can greatly improve the pixel quality and processing speed of images, and also have a certain promoting effect on the improvement of computer technology level. Especially on CUDA based architectures, implementing GPU image processing algorithms through programming languages fully leverages the effectiveness and efficiency of GPU image processing algorithms, which can effectively improve the quality of image or video pixels. At the same time, diversified processing methods also greatly meet the needs of different image processing.

FUND PROJECT

Shaanxi Provincial Education Science 13th Five Year Plan Project (Exploration and Research on the Development Path of Educational Intelligence in Normal Universities under the Background of Artificial Intelligence, No. SGH24H367). Special Research Fund Project of Xianyang Normal University (Research on Text Representation and Classification Methods Based on Deep Learning, No. XSYK24010).

References

[1] Zheng, S., Liu, S., Zhang, Z., Gu, D., Xia, C., Pang, H., & Ampaw, E. M. (2024). Triz method for urban building energy optimization: Gwo-sarima-lstm forecasting model. arXiv preprint arXiv:2410.15283.

[2] Lyu, T., Gu, D., Chen, P., Jiang, Y., Zhang, Z., Pang, H., ... & Dong, Y. (2024). Optimized CNNs for Rapid 3D Point Cloud Object Recognition. arXiv preprint arXiv:2412.02855.

[3] Yin, Y., Xu, G., Xie, Y., Luo, Y., Wei, Z., & Li, Z. (2024). Utilizing Deep Learning for Crystal System Classification in Lithium - Ion Batteries. Journal of Theory and Practice of Engineering Science,?4(03), 199–206. https://doi.org/10.53469/jtpes.2024.04(03).19

[4] Luo, Y., Wei, Z., Xu, G., Li, Z., Xie, Y., & Yin, Y. (2024). Enhancing E-commerce Chatbots with Falcon-7B and 16-bit Full Quantization. Journal of Theory and Practice of Engineering Science, 4(02), 52–57. https://doi.org/10.53469/jtpes.2024.04(02).08

[5] Liu, H., Li, N., Zhao, S., Xue, P., Zhu, C., & He, Y. (2024). The impact of supply chain and digitization on the development of environmental technologies: Unveiling the role of inflation and consumption in G7 nations. Energy Economics, 108165.

[6] Li, T. (2025). Optimization of Clinical Trial Strategies for Anti-HER2 Drugs Based on Bayesian Optimization and Deep Learning.

[7] Jin, X., Yang, Z., Lin, X., Yang, S., Qin, L., & Peng, Y. (2021, April). Fast: Fpga-based subgraph matching on massive graphs. In?2021 IEEE 37th international conference on data engineering (ICDE) (pp. 1452-1463). IEEE.

[8] Yang, Z., Lai, L., Lin, X., Hao, K., & Zhang, W. (2021, June). Huge: An efficient and scalable subgraph enumeration system. In Proceedings of the 2021 international conference on management of data (pp. 2049-2062).

[9] Chen, Y., Tahir, A., Yan, F. Y., & Mittal, R. (2023, December). Octopus: In-Network Content Adaptation to Control Congestion on 5G Links. In 2023 IEEE/ACM Symposium on Edge Computing (SEC) (pp. 199-214). IEEE.

[10] Chen, Y., Yao, R., Hassanieh, H., & Mittal, R. (2023). {Channel-Aware} 5g {RAN} slicing with customizable schedulers. In 20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23) (pp. 1767-1782).

[11] Peng, Q., Zheng, C., & Chen, C. (2024). A Dual-Augmentor Framework for Domain Generalization in 3D Human Pose Estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 2240-2249).

[12] Yan, Q., Yan, J., Zhang, D., Bi, S., Tian, Y., Mubeen, R., & Abbas, J. (2024). Does CEO power affect manufacturing firms' green innovation and organizational performance? A mediational approach. Sustainability, 16(14), 6015.

[13] Ren, F., Ren, C., & Lyu, T. (2025). Iot-based 3d pose estimation and motion optimization for athletes: Application of c3d and openpose. Alexandria Engineering Journal, 115, 210-221.

[14] Fan, Y., Wang, Y., Liu, L., Tang, X., Sun, N., & Yu, Z. (2025). Research on the Online Update Method for Retrieval-Augmented Generation (RAG) Model with Incremental Learning. arXiv preprint arXiv:2501.07063.