

GPU 加速的遥感图像感知哈希认证算法

姜婷婷^{1,2}, 丁凯孟^{2*}

(1. 吉宝通讯(南京)有限公司, 江苏 南京 211189;

2. 金陵科技学院网络与通信工程学院, 江苏 南京 211169)

摘 要: 采用图形处理单元(GPU)对感知哈希算法进行并行化设计与实现。首先,对遥感图像进行预处理,以克服显存大小的限制;然后,基于 GPU 检测遥感图像的特征点;最后,在 CPU 端进行感知哈希序列的生成。实验结果表明,该算法与基于 CPU 的算法相比,计算性能有了显著提高,可满足遥感图像进行认证的需求。

关键词: 感知哈希;遥感图像;认证;GPU 加速

中图分类号: TP751

文献标识码: A

文章编号: 1672-755X(2019)04-0010-05

GPU-accelerated Perceptual Hash for the Authentication of Remote Sensing Image

JIANG Ting-ting^{1,2}, DING Kai-meng^{2*}

(1. Keppel Communication(Nanjing) Co., Ltd, Nanjing 211189, China;

2. Jinling Institute of Technology, Nanjing 211169, China)

Abstract: This paper uses the graphics processing unit (GPU) to implement the parallelization of the perceptual hash algorithm. Firstly, the remote sensing image is pre-processed to overcome the limitation of the GPU memory size. Secondly, the detection of the feature points of the remote sensing image is based on the GPU. Finally, the perceptual hash sequence is generated on the CPU side. The experimental results show that compared with the CPU-based algorithm, this algorithm has significantly improved computing performance and can meet the need of remote sensing image authentication.

Key words: perceptual hash; remote sensing; authentication; GPU-accelerated

随着遥感技术与电子技术的发展,遥感图像的分辨率得到了显著提高,遥感图像在商业及国防科技中的应用也越来越广。但是,在分辨率提高的同时,遥感图像的数据量也迅速增长。例如,海洋遥感数据目前正在以每天 TB 级的速度增长,是公认的“大数据”^[1]。这就对遥感图像的相关处理与应用带来了新的挑战。

另一方面,地理数据安全问题一定程度上关系着国家安全和国防建设^[2]。作为一种重要的地理数据,遥感图像在使用、存储等过程中极易受到各种攻击。一旦遥感图像的完整性、真实性受到影响,那么使用价值将大幅度降低,甚至对国家、国防产生威胁。所以,安全问题对遥感图像的产业化应用形成较大制约。

在遥感图像的多种安全技术中,感知哈希一方面可以识别恶意篡改,另一方面还能对不改变有效内容

收稿日期: 2019-11-23

基金项目: 国家自然科学基金(41801303);江苏省自然科学基金(BK200170116);金陵科技学院基金(jit-b-201520, jit-fhxm-201604);2019 年度江苏省高校“青蓝工程”项目

作者简介: 姜婷婷(1990—),女,内蒙古通辽人,助理工程师,主要从事图像分析等研究。

通信作者: 丁凯孟(1985—),男,山西长治人,副教授,博士,主要从事深度学习、地理数据安全等研究。

的操作具有鲁棒性,可以实现基于感知内容的认证,克服了传统安全技术的不足。但是,随着遥感图像的数据量越来越大,现有感知哈希算法的计算效率逐渐成为突出问题,已经逐渐不能满足遥感图像认证的实时性需求。而随着电子技术的发展,图形处理器(GPU)已成为计算机系统中具备高性能计算能力的部件,已经在许多实时性要求较高的应用领域中获得成功^[3-9]。

目前,GPU已经从单纯的图形运算处理器发展成为具有通用处理能力的处理器,尤其是随着统一计算设备架构(Compute Unified Device Architecture,CUDA)的提出,使得将GPU作为并行计算设备的应用越来越多。本文针对大数据量遥感图像进行感知哈希认证中存在的问题,研究高性能的遥感图像感知哈希认证算法。

1 理论背景

1.1 遥感图像认证与感知哈希

传统的认证技术主要包括Hash函数、数字签名以及脆弱水印技术等。Hash函数与数字签名对数据二进制级别的变化非常敏感,以至于数据只有发生一个比特的变化,就认为数据发生篡改。但是,在遥感图像实际的使用过程中,数据压缩、格式转换、不可见水印嵌入等操作并不改变数据承载的有效内容信息,发生变化的仅仅是信息的载体,所以,Hash函数、数字签名等密码学技术不能完全满足遥感图像的认证问题。脆弱水印对数据进行一定的修改,但这种修改往往会对遥感图像的精度产生影响,而且脆弱水印同样关注数据载体的变化。遥感图像的认证,应当关注遥感图像承载内容是否发生变化,而不是数据载体本身的变化。感知哈希能够克服传统技术的不足,为遥感图像基于感知内容的认证提供了另一种解决方法。

感知哈希实际上是一种基于感知内容的单项映射,能够将具有相同内容的图像、视频等多媒体数据映射为一段数字摘要,并满足鲁棒性和安全性要求^[10]。与Hash函数相比,感知哈希算法最明显的区别在于“鲁棒性”。图像感知哈希具有感知内容、鲁棒性、摘要性、安全性等特点,已经被应用于图像认证^[11-13]、图像检索^[14]、图像拷贝检测^[15]等领域。现有的遥感图像感知哈希认证算法一定程度上满足了需求^[11-12,16-17],但对于数据量较大的遥感图像,计算感知哈希序列的时间较长,无法满足实时性要求较高的应用场合。

1.2 GPU及CUDA架构

近年来,随着电子科技的发展,GPU计算能力的提升越来越明显。GPU在结构上与CPU相比差别很大,天然具有较好的并行性。但是,GPU的主频一般不高,比较适合并行数据处理;而CPU主频较高,能够高效完成串行数据处理。所以,GPU需要与CPU协同工作,不能算是独立运行的计算平台,也可以将GPU视为CPU的协处理器。通常意义上的GPU加速或者GPU并行计算,实际上是指的基于CPU+GPU协同的异构计算架构:GPU与CPU通过PCIe总线连接在一起,CPU一般称为主机端(host),而GPU则称为设备端(device)。

2007年6月,NVIDIA推出专门为GPU硬件设计的并行运算框架的CUDA,提供了GPU编程的简易接口。CUDA采用统一处理架构,而且增加了共享存储器,这些特性使得CUDA更容易发挥并行计算的优势。因此,本文算法利用CUDA实现遥感图像的感知哈希的高性能计算。

2 算法描述

感知哈希的关键步骤在于数据的特征提取,这也是感知哈希算法的计算瓶颈所在,因此,研究高性能的遥感图像感知哈希认证算法,关键在于如何高效地提取遥感图像的特征。本文在现有基于特征点的感知哈希算法的基础上^[18-19],研究基于GPU加速的感知哈希认证算法。如图1所示,本文算法的主要步骤包括:首先,通过GPU并行提取遥感图像的SIFT特征点;接下来,通过SVD对选择的特征点进行降维;最后,通过密码学Hash函数对降维后的特征进行压缩,继而生成感知哈希序列。

算法的具体步骤如下:

步骤1:遥感图像的读取与预处理。由于以下原因,本文算法需要对遥感图像进行预处理,以重置其大小:首先,显存的容量有限,不可能容纳任意大小的遥感图像;其次,为了算法的鲁棒性,本文算法需要对

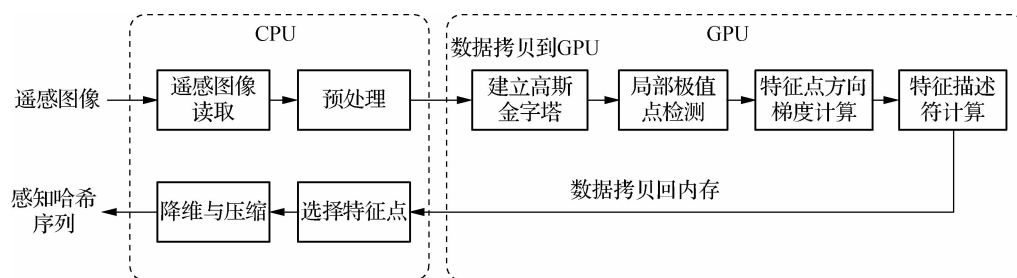


图 1 算法流程

提取的特征点进行选择,数量有限的特征点难以表征较大遥感图像的内容变化。预处理的过程如下:设定阈值 m ,如果遥感图像的长或宽大于 m ,则通过插值算法将其长或宽置为 m 。如果遥感图像的长和宽都大于 m ,则其大小被置为 $m \times m$ 。算法根据显存大小来决定 m 的大小。在本文实验中,根据实验的环境,设置 $m=2\ 048$ 。

步骤 2:基于 GPU 的 SIFT 特征点并行计算。主要包括:建立高斯金字塔、局部极值点检测、特征点方向梯度计算、特征描述符计算。基于 GPU 提取特征点的实现过程,本文算法参考 SiftGPU,限于篇幅,具体流程不再赘述。得到的特征点拷贝到内存,供 CPU 进行下一步处理。

步骤 3:基于 CPU 计算感知哈希序列。主要包括:1)选择特征点。为了增强算法鲁棒性,本文算法从 GPU 计算得到的特征点中均匀选择 128 个特征点。如果特征点的数量不足 128,则用描述符为 0 的“虚拟特征点”进行补齐。2)构建特征矩阵。本算法在选择特征点的基础上构建 128×128 的特征矩阵,其中,每个特征点的 128 维的描述符构成矩阵的一行。3)特征矩阵分解。为了得到遥感图像的鲁棒特征,本文算法对特征矩阵进行 SVD 分解,提取部分较大的奇异值作为原始遥感图像的鲁棒特征。4)特征单向压缩。本算法使用密码学 Hash 函数(实验中以 MD5 为例)对提取的鲁棒特征进行归一化,得到的结果就是遥感图像的感知哈希序列。这样,一方面得到固定长度的序列,另一方面,使算法的安全性有了保障。

遥感图像的认证过程,与现有算法^[11-12,16-17]相似。限于篇幅,这里不再赘述。

3 实验结果与分析

本文选取如图 2(a)~(d)所示的 4 幅遥感图像进行测试与分析,大小分别为 415×396 像素、 $1\ 024 \times 1\ 024$ 像素、 $2\ 048 \times 2\ 048$ 像素、 $3\ 000 \times 3\ 000$ 像素。需要说明的是,由于显存大小的限制,本文算法在预处理阶段对实验图像的大小进行了处理,因此,从测试算法计算性能的角度出发,没有必要过多采用比 $2\ 048 \times 2\ 048$ 像素的图像大很多的图像。

实验硬件平台为:2.8 GHz 主频的 CPU、NVIDIA GTX 1070 GPU、8 GB RAM、内存 16 G;软件平台为 Ubuntu 16.04, CUDA 版本为 9.0,使用 GDAL 的库函数接口实现遥感图像的读取,并基于 OpenCV 的库函数实现部分功能。对比算法采用相同的软硬件平台,并采用 C++ 作为开发语言。



图 2 实验遥感图像

3.1 算法的计算性能测试与对比

本文算法旨在提升计算遥感图像的感知哈希序列的速度,因此,这里将本文算法与纯粹基于 CPU 的

感知哈希算法进行对比。然而,不同感知哈希算法由于针对的目标、应用环境、要求以及实现方式的不同,计算性能上往往有较大差异,将其他类型的算法作为比较对象没有意义。因此,为了尽可能客观地反映本文算法的计算优势,这里,文献[18]设计“CPU 版的本文算法”作为比较对象,也就是将使用 CPU 的感知哈希认证算法作为对比算法。对比算法的流程描述如下:根据遥感图像的大小进行预处理,然后提取 SIFT 特征点,接下来根据特征点的坐标选择一部分特征点构造特征矩阵,最后生成感知哈希序列。经过测试,本文算法与对比算法计算感知哈希序列的如表 1 所示。

表 1 计算性能测试与对比

遥感图像	对比算法用时(基于 CPU)/ms	本文算法用时(基于 GPU 加速)/ms	加速比
遥感图像 A	119	8.6	13.84
遥感图像 B	458	27.6	16.59
遥感图像 C	2 416	85.8	28.16
遥感图像 D	2 613	87.5	29.86

由表 1 可知,本文算法的计算性能有了很大提高,且加速比随着遥感图像的增大而增大。在实际应用中,如果对本文算法的实时性要求更高,可以在预处理阶段设置更小的阈值 m 来进一步提高算法运行效率。另外,本文算法只在特征提取阶段使用 GPU 加速,其他的步骤,如遥感图像预处理、特征矩阵的 SVD 分解等,均在 CPU 上进行计算。

3.2 算法的有效性测试与分析

本文算法虽然以提高计算性能为主要目标,但算法本身的有效性依然不能忽视,否则,即使计算性能提高很多,也将没有意义。本文从以下几个方面测试算法的有效性:

1)篡改敏感性测试。认证过程,需要检测出遭受恶意篡改的遥感图像。如图 3 所示,为实验遥感图像 A 的篡改实例,分别为删除地物、增加地物与变换地物等篡改操作。表 2 为实验遥感图像 A 篡改前后的感知哈希序列变化对比。由表 2 可知,遥感图像经过篡改后,感知哈希序列发生剧烈变化,所以能够有效检测篡改。



图 3 篡改敏感性检测

表 2 篡改前后的感知哈希序列对比

遥感图像	感知哈希序列	篡改检测结果
原始遥感图像	21c2e59531c8710156d34a3c30ac81d5	—
篡改 1(添加地物)	d1457b72c3fb323a2671125aef3eab5d	识别
篡改 2(删除地物)	c3e97dd6e97fb5125688c97f36720cbe	识别
篡改 3(变换地物)	6510c39011c5be704182423e3a695e91	识别

2)鲁棒性分析与测试。本文算法没有像其他面向遥感图像的感知哈希算法^[11-12,17]那样采用隐形网格划分将遥感图像分割为不同的区域,无法采用“网格单元百分百”描述算法鲁棒性,因此,本文以鲁棒性测试来实例说明本文算法保持一定的鲁棒性。这里,分别采用格式转换(原始 TIFF 格式转为 BMP 格式)、数字水印嵌入(最低有效位为例)、数据压缩(以 PNG 压缩为例)来测试算法鲁棒性。实验结果为,

图 2 所示实验数据经过上述操作后,感知哈希序列为发生变化,这就说明,本文算法具有一定鲁棒性。

3)安全性分析。经过分析可以发现,本文算法的安全性依托于所采用的 Hash 函数的安全性。本文从摘要性的角度出发,选择 MD5 为例来论述本文算法,但也可以使用其他安全性更高的 Hash 函数,如 SHA-256,其安全性获得了广泛的认可。所以,算法安全性方面能够满足要求。

4 结 语

遥感图像的数据量越来越大,使得现有基于 CPU 的感知哈希算法的计算性能越来越不能满足实时性的要求。本文针对该问题,研究基于 GPU 加速的遥感图像感知哈希认证算法。实验表明,本文算法的计算性能相对于基于 CPU 的感知哈希算法有了较大的提高。

但是,本文主要从提升计算性能的角度入手,还未及考虑算法的并行策略。因此,下一步的研究方向将是对遥感图像分区域进行基于 GPU 的加速,研究相应的并行策略,并对算法的鲁棒性、稳定性等方面进行深入分析与测试。

参考文献:

- [1] 黄冬梅,耿霞,魏立斐,等. 基于 Henon 映射的加密遥感图像的安全检索方案[J]. 软件学报,2016,27(7):1729-1740
- [2] 朱长青. 地理数据数字水印和加密控制技术研究进展[J]. 测绘学报,2017,46(10):1609-1619
- [3] 曹小鹏,董梁. GPU 加速的交互式医学 CT 图像区域分割[J]. 中国图象图形学报,2016,21(6):815-822
- [4] López-Fandiño J, Quesada-Barriuso P, Heras D B, et al. Efficient ELM-based techniques for the classification of hyperspectral remote sensing images on commodity GPUs[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(6):2884-2893
- [5] 郝昀超,王显珉. 高分辨率遥感图像配准并行加速方法[J]. 计算机科学,2015,42(9):29-32
- [6] Yusuf A, Alawneh S. A survey of GPU implementations for hyperspectral image classification in remote sensing[J]. Canadian Journal of Remote Sensing, 2018, 44(5):532-550
- [7] 孙婷婷,黄皓,王嘉伦,等. 面向 CPU-GPU 异构系统的数据分析负载均衡策略[J]. 计算机工程与科学,2019,41(3):417-423
- [8] Ma Y, Chen L, Liu P, et al. Parallel programming templates for remote sensing image processing on GPU architectures: design and implementation[J]. Computing, 2016, 98(1-2):7-33
- [9] Sánchez S, Ramalho R, Sousa L, et al. Real-time implementation of remotely sensed hyperspectral image unmixing on GPUs[J]. Journal of Real-Time Image Processing, 2015, 10(3):469-483
- [10] 牛夏牧,焦玉华. 感知哈希综述[J]. 电子学报,2008,36(7):1405-1411
- [11] Ding K, Chen S, Meng F. A novel perceptual hash algorithm for multispectral image authentication[J]. Algorithms, 2018, 11(1):6
- [12] Ding K, Yang Z, Wang Y, et al. An improved perceptual hash algorithm based on U-net for the authentication of high-resolution remote sensing image[J]. Applied Sciences, 2019, 9(15):2972
- [13] Zhang Q, Qiao S, Huang Y, et al. A high-performance speech perceptual hashing authentication algorithm based on discrete wavelet transform and measurement matrix[J]. Multimedia Tools and Applications, 2018, 77(16):21653-21669
- [14] Qin C, Chen X, Luo X, et al. Perceptual image hashing via dual-cross pattern encoding and salient structure detection[J]. Information Sciences, 2018, 423:284-302
- [15] Srivastava M, Siddiqui J, Ali M A. A review of hashing based image copy detection techniques[J]. Cybernetics and Information Technologies, 2019, 19(2):3-27
- [16] 丁凯孟,杨晓梅,苏守宝,等. 融合波段特征的多光谱遥感影像感知哈希认证算法研究[J]. 中国图象图形学报,2018, 23(5):708-718
- [17] 丁凯孟,朱长青. 一种用于遥感影像完整性认证的感知哈希算法[J]. 东南大学学报(自然科学版),2014(4):723-727
- [18] 刘兆庆,李琼,刘景瑞,等. 一种基于 SIFT 的图像哈希算法[J]. 仪器仪表学报,2011,32(9):2024-2028
- [19] Vadlamudi L N, Vaddella R P V, Devara V. Robust image hashing using SIFT feature points and DWT approximation coefficients[J]. ICT Express, 2018, 4(3):154-159

(责任编辑:谭彩霞)