

# 超级计算机的图像结构相似度并行处理方法

陈聪梅,都政,井革新,赵晓惠,靳绍巍,陈远磊,饶青雷,邬鸿

(国家超级计算深圳中心(深圳云计算中心),广东深圳518055)

**摘要:**随着图像信息量呈指数级的增长,图像处理面临效率低与质量差的双重挑战,图像结构相似度作为图像质量评价领域常用的一种方法也面临着同样的问题。为此,提出了一种基于超级计算机的数字图像结构相似度并行处理方法。基于超算中心的高性能计算平台实现图像结构相似度的并行处理,采用分块并行计算图像的三种信息,通过与单机计算比较来评价并行处理的性能。实验结果表明,针对数据量大的图像,采用所提出的并行处理方案的耗时远低于单机计算所耗费的时间,可获得较大的加速比,并有效地提高计算效率。可见,所提出的基于超级计算机的图像结构相似度并行处理方法有效可行,为实现图像实时处理提供了可能性。

**关键词:**数字图像;超级计算机;结构相似度;加速比

中图分类号:TP312

文献标识码:A

文章编号:1673-629X(2017)09-0022-04

doi:10.3969/j.issn.1673-629X.2017.09.005

## Research on Image SSIM Paralleling Algorithm with Super Computer

CHEN Cong-mei, DU Zheng, JING Ge-xin, ZHAO Xiao-hui, JIN Shao-wei,

CHEN Yuan-lei, RAO Qing-lei, WU Hong

(National Supercomputing Center in Shenzhen (Shenzhen Cloud Computing Center), Shenzhen 518055, China)

**Abstract:** With the exponential growth of image information, image processing is faced with the dual challenge of low inefficiencies and poor quality. Image Structural Similarity (SSIM), taken as a common method in the field of image quality assessment, is also faced with the identical problem. Therefore, a method of image SSIM parallel processing based on supercomputers is proposed. The image SSIM parallel processing is implemented based on High Performance Supercomputer Center with computing resources. The three kinds of image information are in the parallel calculation by blocking. The performance of the parallel processing is simulated compared with the serial calculation. The experimental results indicate that in allusion to massive image data, it is far lower in the operation time than serial calculation with larger accelerating ratio and higher efficiency, which is effective and feasible, and makes the real-time processing of image possible.

**Key words:** digital image; supercomputing; SSIM; accelerating ratio

## 0 引言

图像是人类获取和交换信息的主要来源,随着计算机技术的不断发展及广泛应用,数字图像作为承载信息的载体在社会生活中扮演着重要角色。然而,随着应用需求的不断扩大,数字图像处理面临着复杂化和实时性的重大挑战。由于人们对图像质量的要求不断提高,从图像中获得的信息越来越多,相应的计算量也越来越大。图像质量评价也成为图像处理的重要环节,图像质量评价方法主要包括主观评价和客观评价<sup>[1-2]</sup>。主观评价方法取决于评价者,难有统一的标准,通常不适用于数字图像处理领域。客观评价方法

通过采用一个或多个评价指标来评价图像质量。

随着图像的数据量越来越大,对图像的任何处理都需要计算机具有强大的计算性能。普通单处理机对图像处理的效率已远不能满足生产对于图像实时处理的要求。由于单机速度提升受制于设备物理的极限,很难满足大数据量的图像处理计算,随着高性能并行处理系统的发展,基于高性能计算的图像并行处理技术在提高图像处理速度上提供了可能性<sup>[3-4]</sup>。高性能计算机是指运行速度非常快的计算机,也称为超级计算机。

为验证图像大规模并行处理的可行性,将国家超

收稿日期:2016-10-28

修回日期:2017-02-17

网络出版时间:2017-07-11

基金项目:国家重点研发计划(2016YFB0201401)

作者简介:陈聪梅(1989-),女,助理工程师,硕士,研究方向为图形图像高性能计算处理;都政,双高级工程师,通讯作者,研究方向为超级计算机、高性能计算、云计算与大数据。

网络出版地址: <http://cnki.net/kcms/detail/61.1450.TP.20170711.1457.082.html>

级计算深圳中心的曙光 6000 超级计算机作为图像结构相似度并行处理的试验研究平台,提出了基于结构相似度的超级计算机图像并行处理方法,为建立基于高性能计算与云计算的数字图像处理平台提供指导性意见与宝贵参考。

1 基础概念

1.1 图像并行处理

数字图像处理是计算机应用领域中一个重要及广泛的分支,随着计算机和集成电路等技术的发展,图像处理无论在算法上、系统结构上还是应用上均取得了长足进步。

图像是连续图像  $f(x,y)$  的一种近似表示,由采样值组成矩阵表示,每一个采样单元为一个像素。并行算法的设计是在有限的空间和时间上处理更多的图像数据,从而提高了图像处理速度。由于图像处理过程中的数据量巨大,且各种算法中包含大量的矩阵乘法运算和卷积运算,为图像处理的并行计算算法设计提供了可能。

图像处理的并行计算的基本思路是借助多个节点协同解决算法问题,把算法分解成各个并行部分,每个并行部分由单独的节点进行计算。对于超级计算机集群计算,一个重要的原则是减少计算过程中节点间的通信,加大计算时间。

1.2 图像结构相似度算法

在图像质量评价领域,往往需要采用客观的评价方法来量化图像质量。传统的图像质量客观评价方法主要有峰值信噪比(Peak Signal to Noise Rate,PSNR)和均方误差(Mean Squared Error,MSE)<sup>[5]</sup>。这两种方法在很多情况下与人的主观感受不符合。因为图像信息高度结构化,并具有纹理特征,图像像素点之间存在很大的关联性,人眼在观察图像时会特别注意图像内容和结构的合理性,因此 Wang Zhou 等提出了图像结构相似度(Structural SIMilarity,SSIM)<sup>[6]</sup>。SSIM 方法通过比较待评价图像与原图像在亮度、对比度及结构信息的差异,得到评价图像质量的综合指标。

SSIM 评价模型图如图 1 所示。

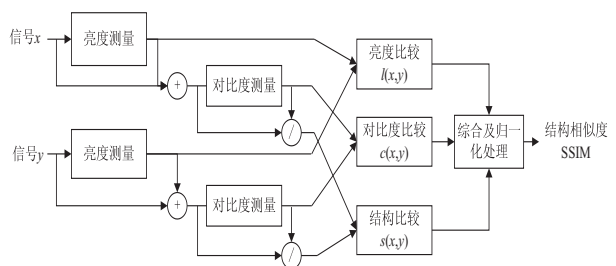


图 1 SSIM 评价模型图

假设  $x$  为参考图像信息,  $y$  表示待评价图像信息。

提取两幅图像的亮度信息  $l$ 、对比度信息  $c$  和结构信息  $s$ ,最后综合三者信息得到图像 SSIM。

(1)亮度信息定义为:

$$l(x,y) = \frac{2u_x u_y + C_1}{u_x^2 + u_y^2 + C_1} \tag{1}$$

其中,  $u_x$  和  $u_y$  分别表示参考图像和待评价图像的均值,计算公式为:

$$u_x = \frac{1}{N} \sum_{i=1}^N x_i \tag{2}$$

$$u_y = \frac{1}{N} \sum_{i=1}^N y_i \tag{3}$$

其中,  $N$  为图像  $x$  的像素数;  $x_i$  为图像  $x$  第  $i$  个像素点的像素值;  $y_i$  为图像  $y$  第  $i$  个像素点的像素值。

(2)对比度信息定义为:

$$c(x,y) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{4}$$

其中,  $\sigma_x$ 、 $\sigma_y$  分别表示图像  $x$  和图像  $y$  的亮度标准差,计算公式为:

$$\sigma_x^2 = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - u_x)^2} \tag{5}$$

$$\sigma_y^2 = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - u_y)^2} \tag{6}$$

(3)结构信息定义为:

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3} \tag{7}$$

其中,  $\sigma_{xy}$  为图像的协方差,计算公式为:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - u_x)(y_i - u_y) \tag{8}$$

为避免分母出现等于零的情况,引入三个常量  $C_1$ 、 $C_2$ 、 $C_3$ 。 $C_1 = (K_1 L)^2$ ,  $C_2 = (K_2 L)^2$ ,  $C_3 = C_2/2$ 。综合式(1)、式(4)和式(7),图像的结构相似度定义为:

$$SSIM\_map(x,y) = [l(x,y)]^\alpha * [c(x,y)]^\beta * [s(x,y)]^\gamma \tag{9}$$

其中,  $\alpha$ 、 $\beta$  和  $\gamma$  表示三者的权重,通常情况下  $\alpha = \beta = \gamma = 1$ 。

1.3 超级计算机基础知识

超级计算机并行系统软件包括超级计算机系统里的高速通信系统、并程序运行支持系统、资源管理系统、并行文件系统等,是硬件系统与应用软件对接的桥梁。国家超级计算深圳中心(深圳云计算中心)部署了由中科院和曙光公司联合研制的曙光 6000 超级计算系统,该系统整体计算能力实测峰值为 1.271 千万亿次浮点运算;操作系统为 64 位 SuSe Linux Enterprise Server 11SP1,采用 IBM Platform LSF 作业管理调度系统;节点间采用 InfiniBand 网络进行通信。

深圳超算中心部署的超级计算机已为广大科研院

校及机构提供了便捷的计算平台,可节约科学研究成本 40% 以上,缩短研发周期一半以上,囊括化学计算、生物计算、材料模拟、电磁学、流体力学计算、气象预报等多个专业领域。在图像处理领域,为提高图像处理速度,有些研究者从事于并行图像处理算法的设计与实验研究,实验表明通过搭建的集群系统设计图像的并行处理算法可提高处理速度<sup>[7-10]</sup>。目前还没有在超算平台进行数字图像处理的应用研究。为此,进行了相关探索,为后续在超算中心建立图像处理云平台打下基础。

2 相似度算法的并行设计与实现

实现并行计算的原理是把一个大计算量的作业分解成多个子任务,分配到各个节点进行并行计算<sup>[11-12]</sup>。结构相似度算法主要针对两幅图像的图像信息进行计算,在对大数据量的遥感图像或红外图像进行质量评价时,采用并行计算,可以极大降低算法运行时间。

2.1 算法并行原理与实现

根据 Wang Zhou 等提出的 SSIM 评价方法,考虑到图像上不同区域的亮度信息、对比度信息和结构信息会随着区域纹理的变化有所区别,因此对图像进行分块后在局部窗口求取 SSIM 的效果优于直接计算整幅图像的 SSIM。分块计算图像的亮度、对比度和结构信息是结构相似度算法中计算量最大的部分,涉及到对图像像素的多项计算,对于数据量大的图像需要耗费过长时间。SSIM 算法并行计算的原理是把计算图像三种信息的任务分解到多个计算节点上并行执行。实现过程为:启动节点读取参考图像与待评价图像,设置参与计算的节点数为 3 的倍数,分别在每个计算节点求取 8×8 局部窗口的均值、标准差和协方差;然后根据式(1)、式(4)和式(7)计算亮度、对比度和结构信息,得到最终的图像平均结构相似度。

2.2 并行处理性能评价标准

评价并行算法性能的指标有多种,主要选用运行时间、加速比,并行效率来说明并行处理的有效性<sup>[13]</sup>,并用实验数据加以验证。

运行时间是指算法在超级计算并行机上从执行到结束的时间。加速比定义为<sup>[14]</sup>:

$$S_p = T_s / T_p$$
(10)

其中,  $T_p$  为在  $p$  个节点上的计算耗时;  $T_s$  为单机运算耗时。

并行效率是指在一个系统上的加速比与该系统上的进程数之比,并行效率定义为<sup>[12]</sup>:

$$E_p = S_p / P$$
(11)

其中,  $P$  为进程数;  $0 < E_p \leq 1$ , 当  $E_p = 1$ , 达到完全

加速比。

2.3 并行计算结果与分析

单机测试环境为 Intel Core i5-2450M 2.5 GHz CPU, 4G 内存; 并行测试环境为曙光 6000 超级计算机。

选取图像像素为 256×256、2 048×2 048 及 5 184×3 456 进行实验, 分别在单机电脑以及超算平台进行计算, 节点数分别为 3、6、9。表 1 为实验得到的测试结果数据。

表 1 实验数据

图像大小	节点数	时间	$S_p$	$E_p / \%$
256×256	单机	0.097 9		
	3	0.063 4	1.544 0	51.47
	6	0.032 4	3.021 6	50.36
	9	0.025 1	3.900 4	43.37
2 048×2 048	单机	1.248 9		
	3	0.641 0	1.948 4	64.95
	6	0.315 8	3.954 7	65.92
	9	0.238 1	5.245 3	58.28
5 184×3 456	单机	7.858 3		
	3	3.529 8	2.226 3	74.21
	6	1.722 1	4.563 2	76.01
	9	1.204 3	6.525 2	72.5

如表 1 所示, 针对实验选取的不同尺寸图像, 采用并行处理的加速比有明显的优势, 验证了图像结构相似度算法采用并行计算的有效性。从加速比和效率可看出, 当选用的节点数增加, 加速比数值越来越大, 效率随节点数的增加呈上升趋势, 但并非无限上升, 当节点数为 6 时达到最大值。为更直观地描述并行处理与单机处理的时间对比, 给出 3 种不同像素尺寸的图像运行时间对比曲线, 如图 2 所示。

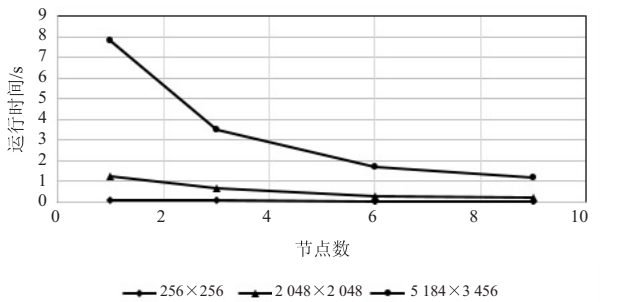


图 2 单机与多节点并行处理时间对比图

结合表 1 的数据, 从图 2 中可知, 对数据量 256×256 的图像, 并行计算的时间与单机比没有太大优势; 而对于大数据量 5 184×3 456 图像的运行时间随着并行节点增加急剧减小, 并逐渐趋于平缓。由此可得出, 对于大数据量的图像, SSIM 算法并行计算具有很大优势, 并行计算的时间明显低于单机处理所用时间。

图3 为并行处理加速比的对比图。

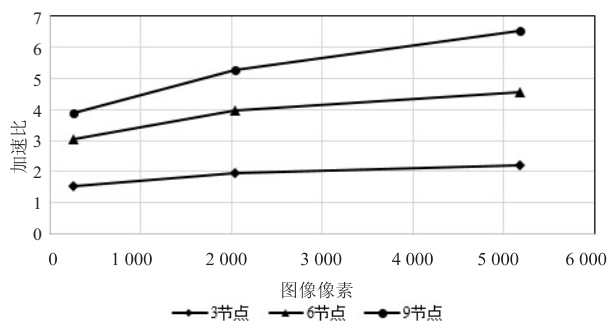


图3 并行处理加速比对比图

从图中可看出,随着数据量的增加,加速比随之增加,表明并行处理速度有所提升,加速比曲线的上升趋势逐渐趋于平缓表明随着并行节点数的提升,处理时间不能无限制上升,当多节点间通讯时间增长到大于图像实际处理时间时,并行效率就无法提升。

图4 为并行效率对比图

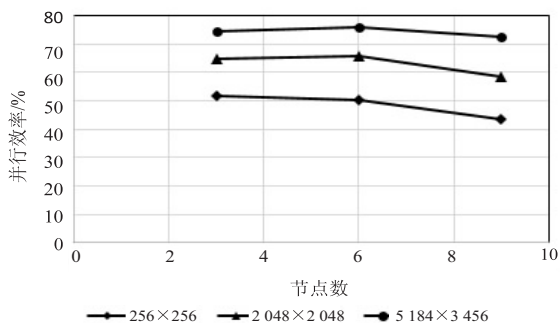


图4 并行效率对比图

由图4可直观看出,随着图像数据规模的增大,效率逐渐增大。当数据量较小时,如 $256 \times 256$ ,使用3节点效率高于更多节点的计算效率;但当数据量越大,多节点并行效率提高,当并行节点为6时,效率达到最大。从而进一步说明,并行处理并非采用节点数越多,并行效率就越高,需要考虑到图像数据在节点间的通信及负载的均衡等。

### 3 结束语

为在超级计算平台实现大数据容量图像并行处理并提升运行速度和效率,提出了一种基于超级计算机的图像结构相似度并行处理方法。该方法采用分块并行计算图像的三种信息,通过与单机计算比较来评价并行处理的性能。实验结果表明,该算法适用于并行处理,对于处理大数据容量的图像,可以有效降低时间

消耗,同时合适的并行节点可有效提高图像处理的计算效率。由分析可知,所设计的SSIM并行处理方法在6节点计算中具有较好的运算效率,且能够随着图像数据量的增大通过增加节点数来提升运算速率;该方法可有效缩短运算时间,为实现图像实时处理提供了可能性,为今后建立基于高性能计算与云计算的数字图像处理平台提供了指导性意见与宝贵参考。

### 参考文献:

- [1] 庞璐璐,李从利,罗 军. 数字图像质量评价技术综述[J]. 航空电子技术,2011,42(2):31-35.
- [2] Xiong Xinghua. Digital image quality assessment[J]. Science of Surveying and Mapping,2004,29(1):68-71.
- [3] 陈国良,毛 睿,蔡 晔. 高性能计算及其相关新兴计算技术[J]. 深圳大学学报:理工版,2015,32(1):25-31.
- [4] 唐少宇. 高性能计算机体系结构综述[J]. 计算机光盘软件与应用,2010(15):51-52.
- [5] 李 航,路 羊,崔慧娟,等. 基于频域的结构相似度的图像质量评价方法[J]. 清华大学学报:自然科学版,2009,49(4):559-562.
- [6] Wang Zhou, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing,2004,13(4):600-612.
- [7] 李金龙. 高性能计算在图像匹配中的应用研究[D]. 西安:西安建筑科技大学,2012.
- [8] 蒋利顺,刘定生. 遥感图像K-Means并行算法研究[J]. 遥感信息,2008(1):27-30.
- [9] Klimeck G, McAuley M, Deen R, et al. Near real-time parallel image processing using cluster computers[C]//International conference on space mission challenges for information technology. Pasadena, CA: [s. n.], 2003:13-16.
- [10] 林 鹏. 基于集群系统的并行图像处理算法的设计与实现[D]. 青岛:中国海洋大学,2006.
- [11] 史凤丽. 基于集群计算机的图像并行处理[D]. 西安:西安科技大学,2010.
- [12] Cwik T, Klimeck G, McAuley M, et al. Applications on high performance cluster computers production of mars panoramic mosaic images[C]//Proceedings of the 2001 AMOS technical conference. [s. l.]: [s. n.], 2001.
- [13] 周益民. 图像处理并行算法的研究[D]. 成都:电子科技大学,2006.
- [14] 金 伟. 基于云计算的数字图像处理平台[D]. 杭州:浙江大学,2011.