

Com4521 / Com6521：使用GPU进行并行计算分

配：第1部分

截止日期：2019年3月18日星期一17:00（第7周）

最后编辑：15/02/2018

标记已分配

作业第1部分（共2部分）占作业总分数的30%。总分配标记（第1部分和第2部分）占模块总标记的80%。

对于代码功能和性能，赋值1标记的权重为50%，通过书面报告证明理解为50%。

文件变更

这是作业文件的第1版。此处将记录对本文档的任何更正或更改，并将更新发送至课程[谷歌群组邮件列表](#)。

介绍

任务的目的是测试您在CPU上实现高效代码的理解和技术能力。您需要对基准和优化简单图像处理问题的实现进行基准测试。您将从实现串行CPU版本开始，然后使用OpenMP将此版本与多核处理器并行化。这项任务的重点在于您逐步改进代码以融合高效实现的能力。为了证明这一点，您需要记录（在书面报告中）为改善性能（并确保正确性）而进行的任何设计考虑。例如，您应该使用基准测试来演示实施的更改如何导致性能改进。除非您在书面报告中表明了对如何逐步完善实施以达到最终解决方案的理解，否则在评估中仅提交具有优异性能的一段代码将无法获得高分。

任务（概述）

任务是在摄影图像输入上实现简单的像素滤镜以产生照片马赛克。马赛克应该平均由输入单元大小（ c ）定义的正方形区域内的像素，该大小应该是2个数的任何正幂（即 $c = 2^n$ ，其中 n 是任何正整数值）。value的值但不得超过图像宽度或高度。图1显示了不同单元尺寸的马赛克效果。

马赛克滤波器的实现细节是镶嵌单元应由镶嵌单元边界内所有像素的平均颜色值组成。因此，对每个单元有贡献的像素数（假设图像宽度和高度是 c 的因子）是 c^2 。马赛克生成没有边界条件，单个像素应该只对单个马赛克单元有贡献。输出图像大小应与输入图像大小完全匹配。同一镶嵌单元格内的任何输出像素应具有相同的RGB颜色值。

除了生成马赛克图像之外，还需要为整个图像生成平均颜色RGB值。对于示例图像，这是rgb值（156, 157, 164）。



图1 - 不同的镶嵌细胞尺寸。左上角显示原始图像，右上角显示 $c = 4$ ，左下角显示 $c = 16$ ，右下角显示 $c = 32$ 。

作业要求（代码）

您应该实现Mosaic任务。您应该从提供的源代码开始，该代码可从以下链接获得。

[链接到起始代码，参考二进制和图像示例](#)

您应该完成TODO部分。您的程序应该接受现有`print_help()`函数描述的参数。该作业具有以下附加要求：

程序输入和输出格式：

程序中输入和输出图像所需的文件格式是二进制和纯文本PPM。可在线获取格式说明（<http://netpbm.sourceforge.net/doc/ppm.html>）。为简单起见，可以假设格式将始终使用换行符来分隔幻数，宽度，高度和最大颜色值。例如，您的代码应以下列格式读写PPM标头

```
P3
#评论标题800中的任何位置
600
255
```

其中P3是幻数，在上述情况下表示纯文本格式（P6表示二进制数据），800是图像宽度，600是图像高度，255是最大颜色值。您的代码只需支持255的最大颜色值（即每个颜色通道8位）。注释可以出现在标题中的任何位置，但不会出现在最大颜色值之后。

对于每个像素值，纯文本PPM的格式应为“rgb \ t”。即每个像素值之后，每个红色，绿色和蓝色值应该用空格分隔。应在每行像素值之后放置换行符（您可以忽略早期格式指南建议的7字符行限制）。测试程序提供了许多样本纯文本和二进制PPM。您可以使用Photoshop或Gimp生成自己的二进制PPM文件。

您需要编写输入和输出函数来处理此文件格式。您的代码应该处理任何大小的图像。

镶嵌计算：

你应该实现马赛克滤镜和图像平均值计算；对于C中的单个CPU内核以及具有OpenMP和C的多核CPU（请注意确保在构建代码时启用OpenMP编译器标志）。镶嵌单元的像素值需要平均多个单元。您应该小心确保有效且正确地处理此问题。

在计算整个图像平均RGB值时，这应该表示每个像素值的平均值，而不是每个马赛克值的平均值。这将确保对于不能被div整除的图像尺寸，将始终获得正确的平均值。

程序时序和命令行输出：

代码的命令行输出应该是应用马赛克过滤器的执行时间和马赛克单元格的平均颜色RGB值。您的代码应该基于程序模式参数以串行方式或OpenMP方式执行（请参阅print_help（）函数）。如果执行模式为ALL，则应向控制台显示每种模式的计时和平均值。在所有情况下都应忽略读写磁盘的定时值。

附加评论：

您应该小心确保所有c值都有效（不包括大于图像宽度或高度的c值）。您的代码也应该处理不是multip的倍数的图像。例如，当c= 32时，应通过具有部分镶嵌单元来处理大小为500的图像。在部分镶嵌单元内，只有有效像素应该有助于马赛克颜色。您必须小心不要超出图像范围（或有效内存）。例如，如果将图1中的原始图像裁剪为500x500，则结果应如图2所示。注意：在对部分镶嵌单元格中的值求平均值时必须小心。

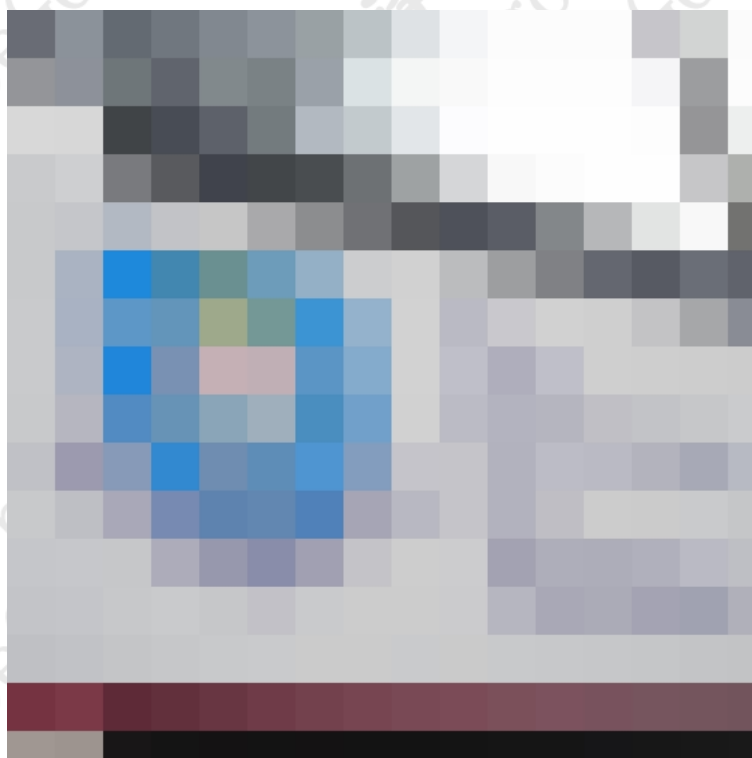


图2：正确处理图像不能完全被马赛克单元格大小整除。注意：右边和底边包含粒子镶嵌值。

测试：

通过验证输出，确保代码为每种模式生成正确的结果非常重要。编译的可执行文件作为赋值传出的一部分提供，它能够在给定输入图像的情况下提供预期的输出。预计您的计划将在标记和评估期间通过一系列盲测。这些测试是盲目的，因为你不知道它们将是什么，除了它们将由不同大小的图像和不同的马赛克单元大小组成。对于每次盲测，都有一个正确的预期输出图像，该图像由参考程序产生。由于可能的舍入错误，在评估代码结果时，可以接受1的余量。

文件要求

您需要记录代码的实现。更具体地说，您需要比较和对比各种技术，以展示您如何融合特定实现。您应该对代码进行基准测试和/或解释为什么一种实现技术（或您所做的优化）比另一种更好。应考虑以下因素：

- 用于存储像素和马赛克值的数据结构的选择。
- 实施方法的比较。在计算镶嵌细胞值时，最好是在像素或镶嵌细胞上循环吗？
- 用于马赛克计算的OpenMP并行化策略的比较。并行内部或外部循环（它们存在于代码中）是否更有效？
- OpenMP技术的比较，可用于计算平均值。你如何避免竞争条件？

- 您为改进马赛克计算循环所做的任何优化。这对性能有何影响？
- 您已应用于代码的串行或OpenMP版本的实现或优化技术的任何其他有趣方面。

应该在各种大的（至少2048×2048）图像大小和values值上进行良好的基准测试。对于代码的每个重大改进，尝试在更改之前和之后显示代码的性能。您应该突出（使用短代码示例）您所做的任何新颖方面或优化。

项目交接

您应该通过MOLE将您的程序代码与单个zip文件中的单个pdf一起提交。您还应该包括visual studio解决方案（或makefile）和任何项目文件。您的代码应该在Visual Studio发布模式配置中构建，而不会出现错误或警告（或者对于具有相应Makefile的Linux）。即使您尚未完成整个作业，也应提交已完成的任何内容。您的代码不应该依赖任何第三方库或工具。

如果您使用任何实验室解决方案中未经修改的代码，您应该明确说明代码不是您自己的注释代码。

印记

转让第1部分的标记将分发如下：

- 50%的任务是针对编码方面的。
 - 50%的编码标记用于编程的质量和代码的性能。
 - 50%的编码标记用于满足要求。
- 50%的任务是用于生成描述您为实现和优化代码而采取的流程的文档。这应该包括文档要求中描述的方法的基准测试和迭代改进。

在评估您的工作时，将考虑代码方面的以下要求。

1. 代码在功能上是否正确？即模拟是否已正确实施并产生正确的行为？
2. 迭代改进代码是否为每个CPU和OpenMP实现产生了足够优化的最终程序？
3. 代码是否充分利用了内存带宽？
4. OpenMP实现是否避免了竞争条件？
5. OpenMP变量是否正确作用域？他们是否过度/不必要地使用共享变量？
6. 是否有任何编译器警告或危险的内存访问（超出分配的内存范围）？你的程序是否释放了所有分配的内存？
7. 您对输入文件的处理是否正确且健壮？程序是否优雅地处理错误的输入文件格式化（例如，引发错误并退出而不是崩溃）？
8. 您的代码是否结构清晰且评论良好？

在评估您的文档时，将考虑以下内容，并作为讨论代码增量改进的指南。

1. 技术描述及其实现方式。

2. 您的文档是否描述了对代码的优化并显示了这些优化的影响？
3. 是否已考虑使用一系列OpenMP方法来并行化问题？对调度进行了适当的调查？对基准测试结果有很好的解释吗？
4. 是否已经探索了一系列OpenMP方法来实现马赛克和平均图像RGB值的计算？对于为什么选择特定方法有很好的解释吗？

开发代码和文档的技巧

您应该从简单的串行实现开始，并描述如何迭代地提高文档中的性能。如果您无法完成某些特定部分（例如，使用OpenMP的活动地图），那么您应该默认使用该部分的简单CPU版本，以便您的代码正确构建并执行生成正确的结果。同样，如果您应用的技术不能提高性能，则应将其包含在文档中，并解释您对其未能按预期工作的原因的信念/理解。您可以使用#define来允许您的代码以不同的版本构建，以便更直接地比较技术。

您应该评论您的代码，以明确您所做的事情。您应该测试您的代码，以确保它适用于所有图像大小，values值和图像输入文件。对于不正确的值和输入文件（例如c= -10或具有不正确像素值数的输入文件），您的代码应优雅地退出并显示有用的错误消息。您的代码永远不应该读取或写入分配的内存。

作业帮助

一般实验课程将提供您的作业帮助。对于实验室之外的具体问题，您应该使用Google讨论组课程。

反馈

您将在复活节假日后收到有关您的作业的反馈，您在交付作业第2部分时应该考虑这些反馈。