

函数式语言程序设计 兴趣作业

问题

我们需要解决一个图像匹配问题。该问题的定义如下：

给定图像 P 和 P' ，求 P' 最有可能出现在 P 的位置。

举个例子，如图 1 所示，左侧的图为原始图像 P ，右侧为 P' ，它是 P 的一个局部图像。如果我们规定图像的坐标系原点为左上角顶点， x 轴水平向右延伸而 y 轴垂直向下延伸，那么可以求得 P' 的左上角顶点在 P 中最有可能的位置是(881,237)。

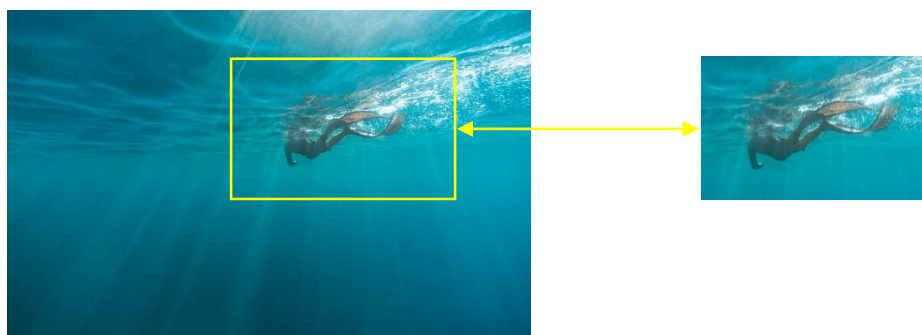


图 1：图像匹配问题示意图

然而 P' 并非一定和原图的局部完全一致。在很多应用场景下， P' 存在一些干扰性的细节，比如模糊，颜色增强，椒盐噪声，污染等，如图 2 所示。因此，高效且准确的图像匹配是一个充满挑战的工作。



图 2：几种图像干扰

为了简化问题，我们给出以下几个假设：

- 1) P' 来自 P 的局部，且不会发生几何变换，如缩放、旋转等。也就是说， P' 一定能够和 P 的某个局部一一对应。因此，我们用 P' 左上角顶点在 P 中的坐标来表达 P' 在 P 中出现的位置。
- 2) P' 只存在如下几种干扰：(a) 基于线性平滑滤波器的网格模糊；(b) 颜色增强，包括色彩、饱和度、亮度的改变；(c) 椒盐噪声(均匀分布的黑白噪声)；(d) 黑色污染。

要求

用 Haskell 实现一个图像匹配的程序。要求如下：

- 1) 作业程序应当能够用 GHC 编译为可执行文件，因此请确保程序中存在 main 函数。
(具体的 GHC 使用方法请参照相关文档)
- 2) 假设可执行文件为 match，那么它的使用方法应该如下

```
match [input image 1] [input image 2] [mode]
```

前两个参数分别对应图像 P 和 P' ，第三个参数是可选项，代表 match 的工作模式。如果你的程序需要针对不同的干扰指定不同的工作模式，请遵循下面的规约：

mode	干扰
0	无干扰
1	线性平滑模糊
2	颜色增强
3	椒盐噪声
4	黑色污染

测试用例包含原图，子图及其干扰模式，因此不用担心需要自己识别干扰模式。

- 3) 可执行文件的输出是一个坐标，即子图的左上角顶点在原图中的坐标。原图中坐标系的规定如图 3 所示。原图的左上角顶点为(0,0).



图 3：图像坐标系

- 4) 作业程序需要支持无干扰的图像匹配和至少 2 种干扰模式的图像匹配(实现得更多会有加分)。请在文档中注明自己的程序支持的干扰模式。
- 5) 输入的图像为 24 位 BMP 文件格式。因此，作业程序需要支持解析这种格式。

贴士

- 1) 你的程序首先需要解析 24 位 BMP 图像文件。由于图像并非文本文件，所以请用相应的二进制文件操作。BMP 的格式规律性较强，请参照相关文档(比如：https://en.wikipedia.org/wiki/BMP_file_format)正确解析 BMP 文件内容。

- 2) 图像匹配问题本质上是一个优化问题。我们可以用一个函数 $\Phi(P', x, y)$ 来表达 P' 和坐标为 (x, y) 的 P 的子图的匹配程度，该函数的取值越大则匹配程度越高。那么，这个优化问题可以表达为

$$\operatorname{argmax}_{(x,y)} \Phi(P', x, y)$$

解决这个优化问题的难点有二：(a) 匹配度函数如何定义；(b) 如何求解最优解。

第一个问题，匹配度函数可以是像素颜色层次的，比如两幅图像对应像素颜色的相关系数：

$$\rho(x, y) = \frac{\sum_{i=0}^{w-1} \sum_{j=0}^{h-1} (C(x+i, y+j) - \overline{C_{(x,y)}})(C'(i, j) - \overline{C'})}{\sqrt{\sum_{i=0}^{w-1} \sum_{j=0}^{h-1} (C(x+i, y+j) - \overline{C_{(x,y)}})^2} \sqrt{\sum_{i=0}^{w-1} \sum_{j=0}^{h-1} (C'(i, j) - \overline{C'})^2}}$$

其中 $\overline{C_{(x,y)}}$ 和 $\overline{C'}$ 分别为 P 的相应区域和 P' 的颜色均值。

匹配度函数还可以是图像特征层次的，这种情况下我们需要抽取两个图像的特征向量再对特征向量进行匹配。

第二个问题，最优解的求解方法本质上是搜索，因此可以使用遍历或带有策略的搜索方法(导向搜索，随机游走等)求解最优解。

- 3) 测试用例中可能会包含大图像，正确使用 Haskell 的并行特性有助于提高程序的性能。(我们的测试机器有 4-8 个物理核心)

评价标准

两个指标：(1) 程序计算出的坐标值与正确坐标值的差距，越小越好；(2) 程序的执行时间，越少越好。两项指标的权重视大家作业的完成情况来确定，但正确性(即第一项指标)不应低于 60%。

我们会使用单独一套测试集来评估作业程序。

作业提交

- 1) 作业以 zip 压缩包格式提交(请使用 WinRAR, WinZip, 7zip, tar 等工具压缩，勿使用快压、好压等工具)，文件名格式为：[学号]_[姓名]_兴趣作业.zip. 压缩包内的组织方式如图 4 所示。

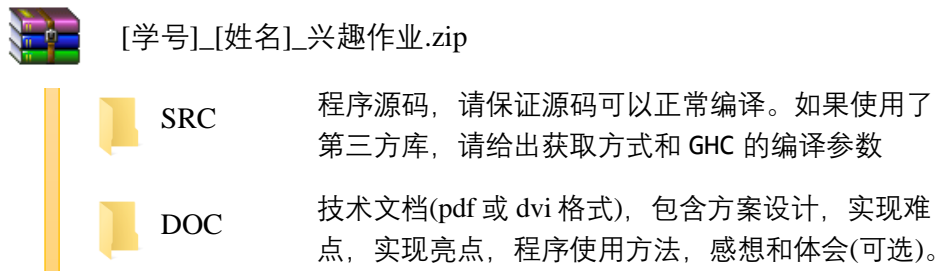


图 4：作业文件的组织方式

- 2) 请独立完成作业，严禁抄袭。

- 3) 作业的截至时间为第 18 周周日晚 23:59。请按时提交作业，否则会影响你成绩的登入。