北京工业大学

**《数字图像处理》实 验 报 告（八）**

|  |  |  |  |
| --- | --- | --- | --- |
| 课程名称 | 数字图像处理 | 实验名称 | 实验八 |

|  |  |  |  |
| --- | --- | --- | --- |
| 学号 | 21071003 | 姓名 | 高立扬 |

|  |  |  |  |
| --- | --- | --- | --- |
| 班号 | 210710 | 实验日期 | 2023.11 |

评阅意见：

成绩 教师签字：

年 月 日

1. 背景知识
2. OTSU（大津法）

OTSU，即最大类间阈值法，又称为大津法，由日本学者大津展之发明。该算法的基本思想是：用一个阈值(threshold)将图像中的数据分为两类，第一类图像的像素点灰度小于这个阈值，第二类反之，这样就得到了图像的背景和前景，或者说A、B类。而如果背景和前景两类图像之的类间方差越大，就说明当前获取到的阈值是最佳的。一般来说，我们希望通过OTSU分割出图像的前景。

大津法的公式是：

其中灰度级k为需要求解的阈值，为整个图像的灰度值，是灰度级k的累加均值，是在阈值为k时，像素被分到A类的概率，为A类的方差。

在MATLAB中，greythresh函数采用的就是大津法，调用它即可获取上文提到的最佳阈值，配合im2grey函数，传入原图和阈值即可实现大津法对前景和背景的二值化分割。

1. 连通分量分析——MATLAB中的bwconncomp函数

本函数查找二值图像中的连通分量并计数，对于二维图像可以设置连通性为4连通或8连通，默认为8连通。函数会返回一个结构体，包含了：①连通性②输入数组的大小③连通分量数量④一个元胞数组，包含了每一个连通分量所包含像素的线性索引向量

1. 个人理解

在数字图像处理领域，二分问题通常需要先将图片转化为二值图。将图像转换为二值图有很多种办法比如双峰法、P参数法、迭代法和OTSU法等。OTSU于1979年被大津展之发明，一直到今天都是十分好用的二值化方法，其优点是计算简单快速，不受图像亮度和对比度的影响，但是当图像噪声敏感，只针对单一目标分割，或者目标和背景大小比例悬殊的时候，效果就会欠佳。而本实验的染色体电镜图，存在背景噪声极多、部分染色体不清晰的问题，用大津法直接分割效果会很差，因此还需要借助连通分量分析算法，对二值图进行进一步的优化，过滤掉连通的噪点，最后实现染色体计数。

1. 实验思路
2. 实验设计
3. 首先读取图片，并用OTSU进行二值化
4. 反转黑白像素，使得前景颜色为白色，并多份拷贝，便于进一步的实验。
5. 设置连通性为4连通，设置一个列表，用来寻找过滤噪声的最佳连通像素数，定义其为过滤参数，过滤后再次进行连通性分析，得到染色体的计数。
6. 设置连通性为8连通，重复上述实验的步骤
7. 对比4连通和8连通哪一个更好，提出进一步的优化策略
8. 得出结论
9. 实验环境

MATLAB R2022a - academic

1. 实验过程
2. 首先手动计数图中染色体数目。由于缺乏专业知识，对于图中一处交错的染色体无法分辨，初步计数为45条，联想到人的染色体数目为46条，因此断定此图为人体细胞染色体电镜图，所以计数为46条
3. OTSU分割原始图像，并反转，得到OTSU.png和reverse.png如下图1-2.所示，图3.为原图，作为对比。可以发现，个别染色体出现了“切成两半”的情况，但是从原图可以看出，“切成两半”的染色体，原本就因为电镜扫描不清楚所导致。因此初步分析认为，连通性为4连通和8连通的连通分量分析，对该图像的染色体计数结果将会不同，下面实验时应当分别进行实验分析。

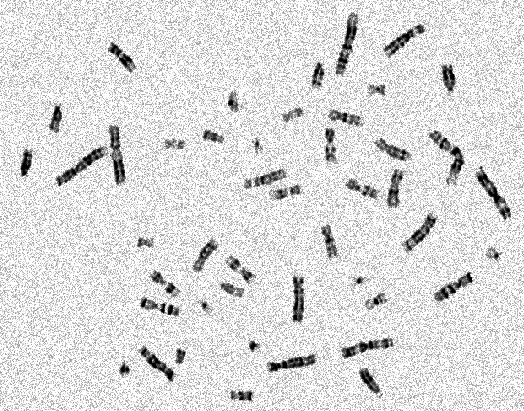
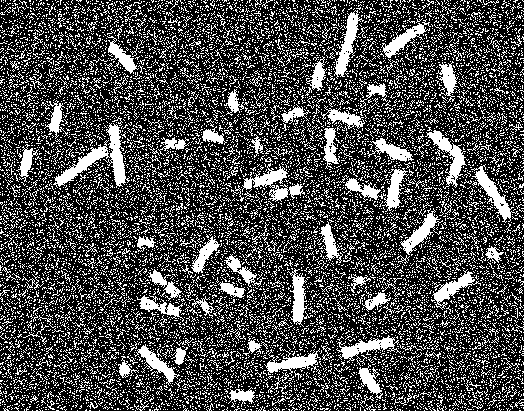
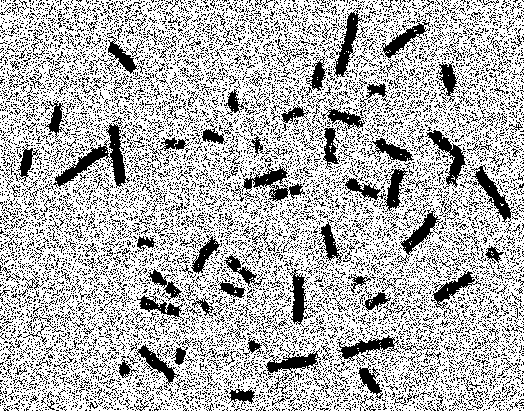


图1.OTSU.png 图2.reverse.png 图3.原图

1. 如下图4.代码所示，连通性测试及噪声过滤参数寻优过程如下：首先进行连通分量分析，并进行连通分量像素计数，经过试验后设置过滤参数列表为[10, 15, 20, 25, 30, 35, 40, 45, 50],通过循环对连通噪声进行过滤，最后分别用两种连通性进行计数，进行分析。

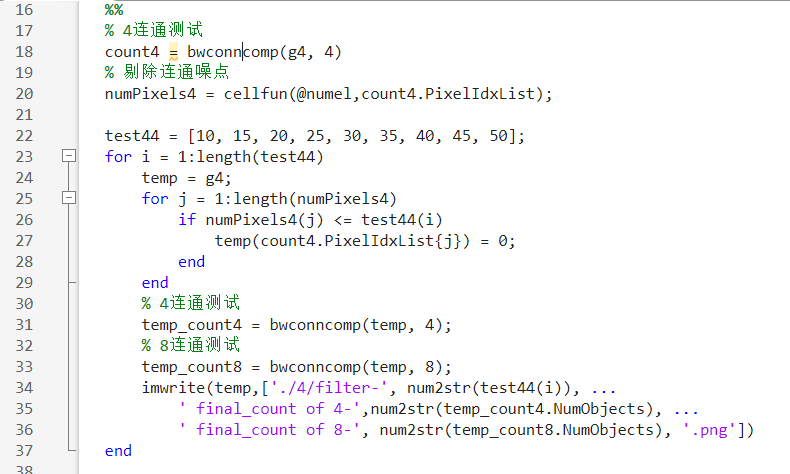


图4.测试代码

全部的实验结果见文件夹“4”。如下图5-6.和表1所示，过滤参数不同导致了部分噪声无法去除，经实验表明，当过滤参数为25时，所有噪声均已去除。



图5-6.左为参数10 右为参数25

表1.连通性为4连通

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 过滤参数 | 4连通计数 | 8连通计数 |  | 过滤参数 | 4连通计数 | 8连通计数 |
| 10 | 145 | 137 |  | 35 | 48 | 47 |
| 15 | 60 | 57 |  | 40 | 48 | 47 |
| 20 | 50 | 49 |  | 45 | 48 | 47 |
| 25 | 48 | 47 |  | 50 | 48 | 47 |
| 30 | 48 | 47 |  |  |  |  |

可见，8连通计数更加准确一些，但是计数结果为47条，多了一条，不过观察“切成两半”的染色体可以得知，无论4连通还是8连通都会将这“两半”染色体计数为2条，因此认为此时计数正确。从结果还可以看出，4连通过滤参数波动较大的情况下，计数情况却很稳定。

那么8连通会不会效果更好一些呢？请看接下来的实验

1. 重复3.的步骤，经过试验后将过滤参数列表更改为[30, 40, 50, 55, 75, 100, 113 ,125]。全部的实验结果见文件夹“8”。如下面图7-9.所示，结合实验结果表明最佳过滤参数在100-113之间

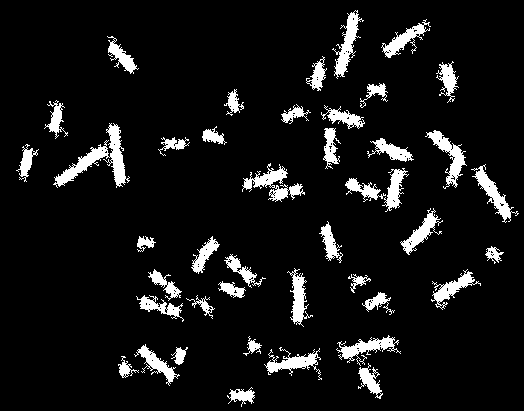


图7-9.左55 中113 右125

显而易见，8连通过滤噪声后，染色体的“毛刺”很多，因此如表2.所示，4连通分析的结果不准确，而8连通计数会将紧挨着的两条染色体算为一条，并且还凸显出了其计数误差大的缺点。当过滤参数为55时，虽然计数正确，但是由图7.明显看出此时有噪声还没有过滤，而参数为50时有更多处未过滤，详见文件夹。

表1.连通性为8连通

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 过滤参数 | 4连通计数 | 8连通计数 |  | 过滤参数 | 4连通计数 | 8连通计数 |
| 30 | 1723 | 82 |  | 75 | 1218 | 45 |
| 40 | 1420 | 57 |  | 100 | 1179 | 44 |
| 50 | 1280 | 48 |  | 113 | 1179 | 44 |
| 55 | 1239 | 46 |  | 125 | 1169 | 43 |

那么用4连通来优化8连通的结果，如何？请看下面实验

1. 首先采用8连通、过滤参数为113来过滤噪声，之后用4连通、过滤参数为25来修理“毛刺”，得到根目录下的图片“repair-4-48-8-47.png”对比结果如图10-12.所示，发现效果甚微，只有及其微小的几个像素被修理，结果为4连通计数48,8连通计数47。

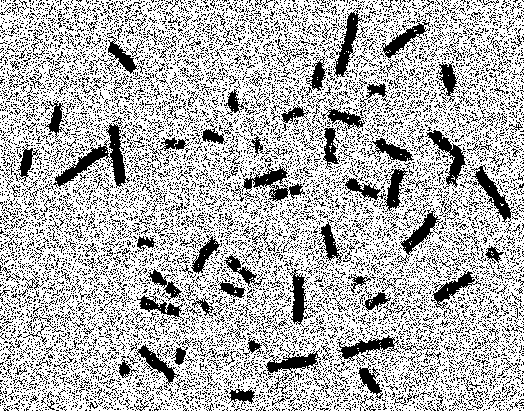
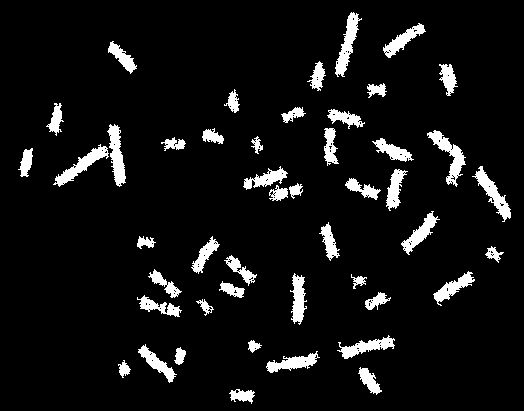


图10-12.左为repair 中为OTSU 右为4连通、参数25

1. 个人认为①在染色体计数问题上，用采用4连通、过滤参数为25来过滤噪声，用8连通来计数最为妥当，并且计数波动小。②一张图片偶然性太大，4连通来修理8连通“毛刺”的作用并不会收效甚微，这在染色体交错折叠或临近的情况下将会大有所用，个人认为可行。
2. 实验总结

实验结果表明

①OTSU法对于本实验这种底噪大，目标不清晰的情况，效果欠佳

②过滤底噪适合采用先8连通后4连通的方法

③计数染色体适合采用8连通

④4连通的过滤参数推荐为25,8连通的过滤参数推荐为113

1. 感想

第一眼看到这个实验，我就想到了灰度图二值化和连通分量分析，看了眼指导书果然如此，说明我本科两年多的自学没有白费。

做实验之前，我对二值化的理解只是一行函数调用，通过本次实验的学习，我才了解到二值化原来分为多种方法，是一个研究至今的科学问题，同时也对近几十年来不断研发二值化算法的，坚持不懈的科学家们心生敬佩。本次实验中，我大大简化了代码书写，去除了不必要的步骤，使得其一目明了。在观察OTSU分割图，初次使用连通分量分析后，我通过自主思考想到了噪点过滤，并将想法付诸于实践，完成了本实验，取得了完全符合预期的结果，这种结合所学并超越所学的感觉令我十分激动。