

**MyBox：简易工具箱
用户手册 – PDF 工具**

作者：Mara

版本：5.0

日期：2019-4-16

内容目录

1 资源地址.....	3
2 文档.....	3
3 PDF 工具的菜单.....	4
4 查看 PDF-图像模式.....	5
5 提取 PDF 中的图片.....	8
5.1 源 PDF.....	8
5.2 查看 PDF 的信息.....	8
5.3 目标文件.....	9
5.4 预览.....	10
5.5 进程控制.....	11
6 提取 PDF 中的图片-批量.....	12
7 提取 PDF 中的文字.....	13
7.1 源 PDF.....	13
7.2 目标文件.....	13
7.3 预览.....	13
8 提取 PDF 中的文字-批量.....	14
9 将 PDF 页面转换为图像.....	15
9.1 源 PDF.....	15
9.2 目标文件.....	16
9.3 预览.....	17
10 将 PDF 转换为图片-批量.....	18
11 将多个图片合成 PDF.....	19
11.1 管理源图片的列表.....	19
11.2 设置 PDF 的页面尺寸.....	20
11.3 设置图片压缩格式.....	20
11.4 设置页边、作者、页眉.....	21
12 压缩 PDF 的图片.....	22
13 合并 PDF.....	23
14 分割 PDF.....	24

1 资源地址

这是利用 JavaFx 开发的图形化界面程序，目标是提供简单易用的功能，免费开源。项目主页：

<https://github.com/Mararsh/MyBox>

每个版本的源代码、编译好的包、和文档都在 Release 目录下：

<https://github.com/Mararsh/MyBox/releases>

欢迎在线提交软件需求和问题报告：

<https://github.com/Mararsh/MyBox/issues>

云盘地址：

https://pan.baidu.com/s/1fWMRzym_jh075OCX0D8y8A#list/path=%2F

2 文档

本文档介绍 MyBox 的 PDF 工具，下载地址：

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-PdfTools-zh.pdf>

其它文档：

《MyBox 用户手册-综述》

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-Overview-zh.pdf>

《MyBox 用户手册-图像工具》

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-ImageTools-zh.pdf>

《MyBox 用户手册-桌面工具》

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-DesktopTools-zh.pdf>

《MyBox 用户手册-网络工具》

<https://github.com/Mararsh/MyBox/releases/download/v5.0/MyBox-UserGuide-5.0-NetworkTools-zh.pdf>

3 PDF 工具的菜单

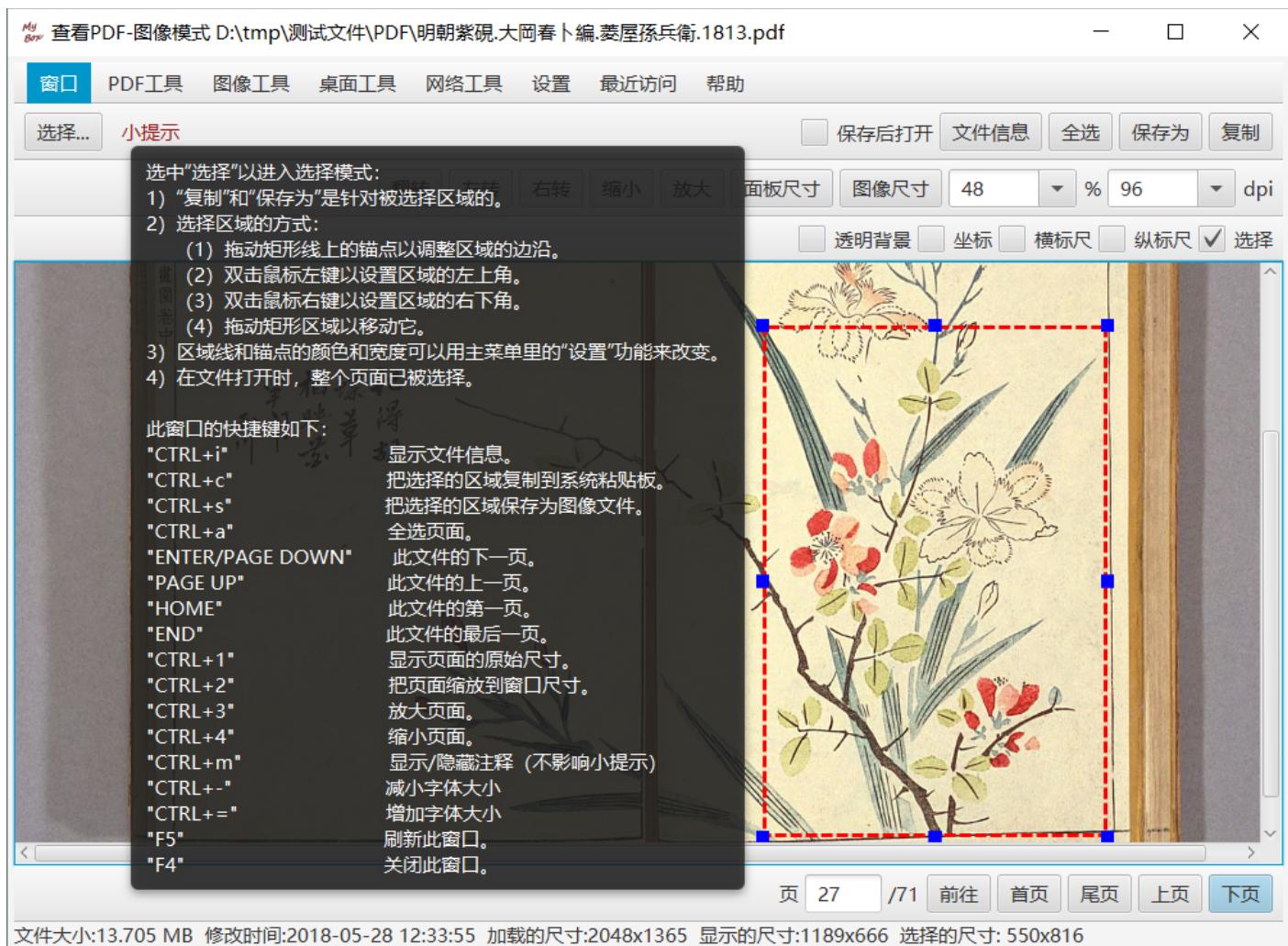


4 查看 PDF-图像模式

1. PDF 文件的每一页被转换为图片逐页显示。
2. 可以在页面图片上选择需要剪切的矩形范围：拖动锚点以调整区域，或者左键取左上角、右键取右下角。
3. 可以把整个页面图像或选择的矩形范围复制到系统剪切板（CTRL+c）。
4. 可以把整个页面图像或选择的矩形范围保存为图像文件（CTRL+s）。
5. 可以选择图像背景是否透明。这个选项只对文字有效。
6. 可以缩放图片的显示（并不缩放图片本身）。
7. 可以旋转图片，保存时也生效。
8. 可以设置图片的密度（dpi）。密度越大，图像越清晰、图像文件越大。
9. 选择是否显示坐标和标尺。

此界面快捷键包括：

"CTRL+i"	显示文件信息。
"CTRL+c"	把选择的区域复制到系统粘贴板。
"CTRL+s"	把选择的区域保存为图像文件。
"CTRL+a"	全选页面。
"ENTER/PAGE DOWN"	此文件的下一页。
"PAGE UP"	此文件的上一页。
"HOME"	此文件的第一页。
"END"	此文件的最后一页。
"CTRL+1"	显示页面的原始尺寸。
"CTRL+2"	把页面缩放到窗口尺寸。
"CTRL+3"	放大页面。
"CTRL+4"	缩小页面。
"CTRL+m"	显示/隐藏注释（不影响小提示）
"CTRL+-"	减小字体大小
"CTRL+="	增加字体大小
"F5"	刷新此窗口。
"F4"	关闭此窗口。



MyBox 查看PDF-图像模式 D:\玛瑞\Mybox\文档\MyBox-UserGuide-5.0-Overview-zh.pdf

窗口 PDF工具 图像工具 桌面工具 网络工具 设置 最近访问 帮助

选择... 小提示 保存后打开 文件信息 全选 保存为 复制

翻转 左转 右转 缩小 放大 面板尺寸 图像尺寸 100 % 96 dpi

474,151 透明背景 坐标 横标尺 纵标尺 选择

100 200 300 400 500 600 700

MyBox 用户手册 - 综述 v5.0

100 5 当前版本
当前是版本 5.0，以下简要介绍已实现的特点。
474,151

200 5.1 跨平台
纯 Java 实现且基于开源库，因此 MyBox 可运行于所有支持 Java 的平台。

300 5.2 国际化
所有代码均国际化。可实时切换语言。目前支持中文、英文。扩展语言只需编辑资源文件。

400 5.3 PDF 工具
1. 以图像模式查看 PDF 文件，可以设置 dpi 以调整清晰度，可以把页面剪切保存为图片。
2. 将 PDF 文件的每页转换为一张图片，包含图像密度、色彩、格式、压缩、质量、色彩转换等选项。
3. 将多个图片合成 PDF 文件，可以设置压缩选项

页 9 /51 前往 首页 尾页 上页 下页

文件大小:7.082 MB 修改时间:2019-04-16 10:36:36 加载的尺寸:793x1122 显示的尺寸:924x466

5 提取 PDF 中的图片

5.1 源 PDF

可以填写要处理的页码范围，若打开 PDF 需要密码则应当填写密码。



5.2 查看 PDF 的信息



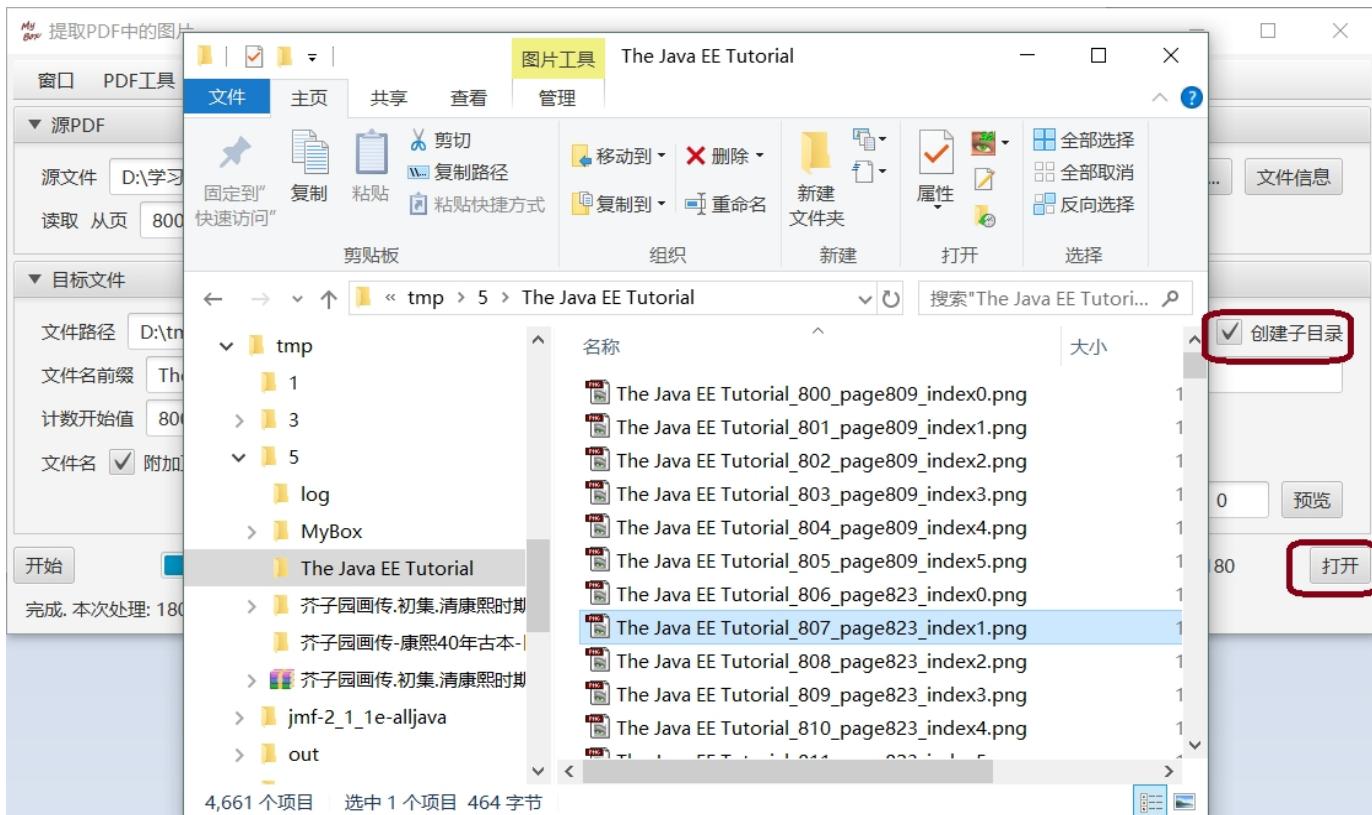
5.3 目标文件

若选择“创建子目录”，则程序将在目标目录中以 PDF 文件名创建一个子目录，将提取出来的图片放入此子目录。若不选此项，则提取出的图片直接放在目标目录下。

提取的图片文件名是：用户填写的“文件名前缀”（被自动初始化为 PDF 文件名）+ 计数值 + 图片扩展名。用户可以设置计数开始值。注意：由于 PDF 页可能有多个图片，因此图片计数值不等于页码。

用户可以选择文件名是否附加页码或者页内图片序号

点击底部的“打开”按钮，可以查看图片在目录文件的生成情况：



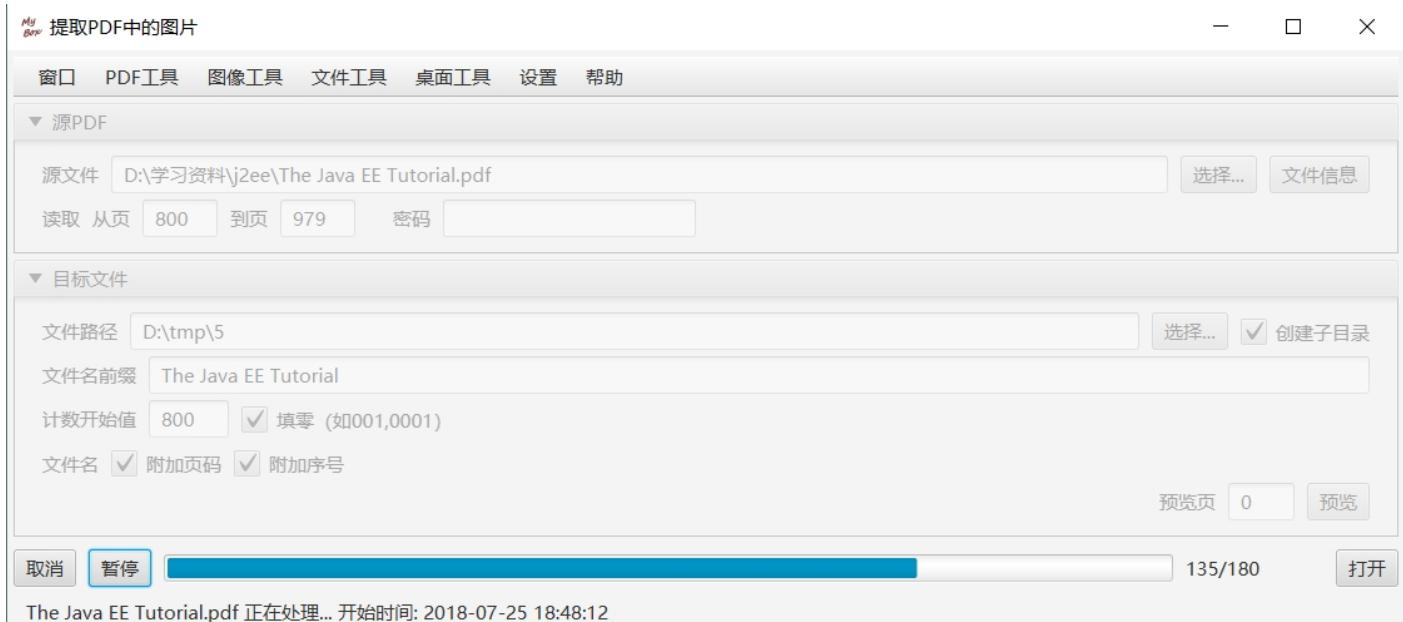
5.4 预览

填写“预览页”、点击“预览”按钮，则程序提取预览页的第一图片后显示此图片。若填写的页码内没有图片，则会弹出“没有数据”的提示，用户可以另选预览页。



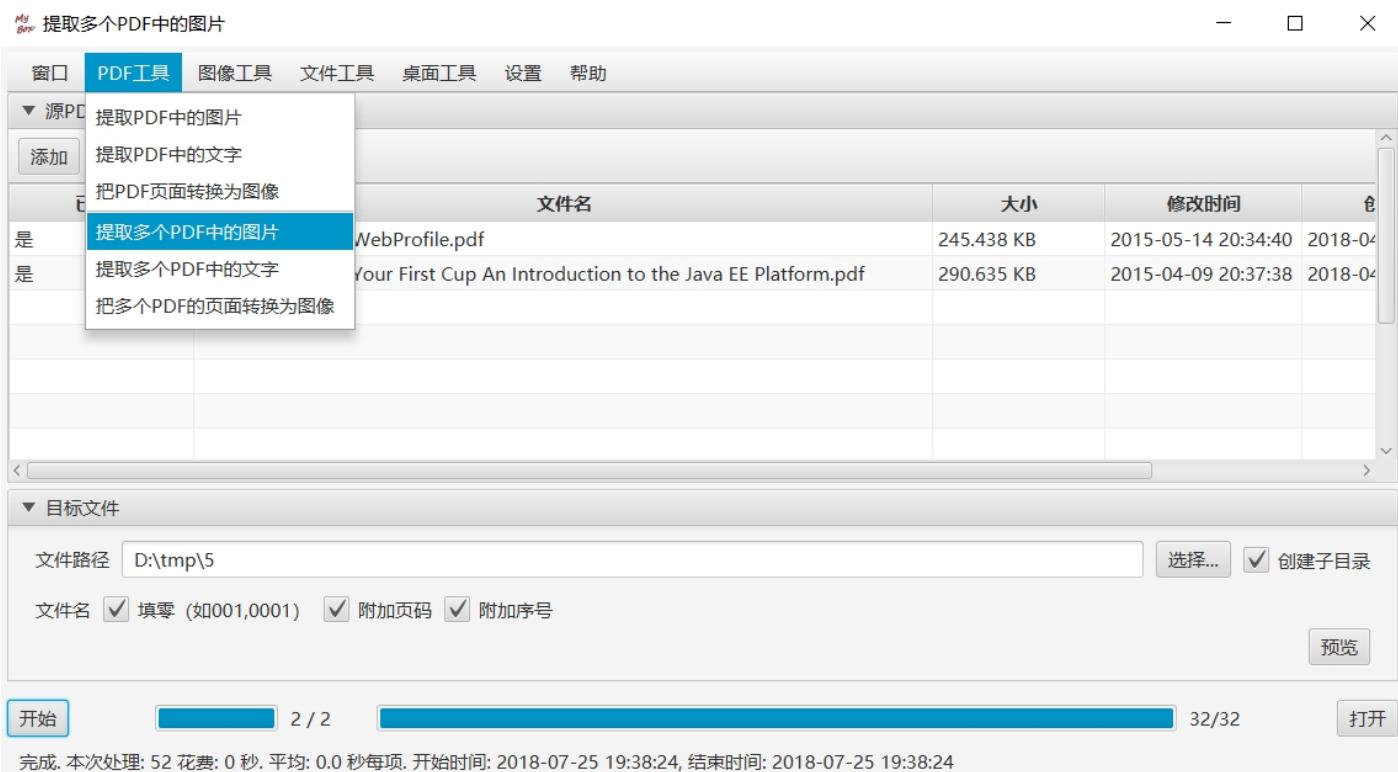
5.5 进程控制

点击底部“开始”按钮则开始提取进程，用户可以点击“暂停”/“继续”/“取消”来控制进程。



6 提取 PDF 中的图片-批量

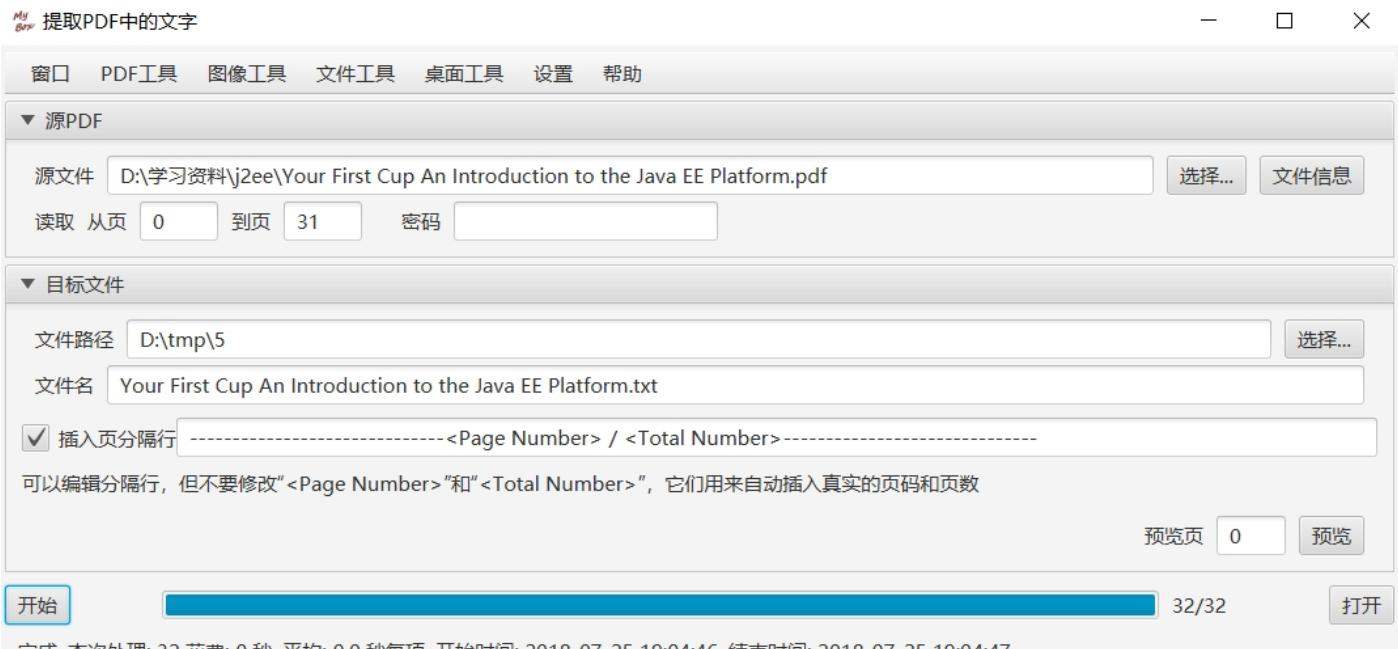
可以对多个 PDF 文件执行批量提取图片：



7 提取 PDF 中的文字

7.1 源 PDF

可以填写要处理的页码范围，若打开 PDF 需要密码则应当填写密码。点击“文件信息”按钮，可以查看 PDF 的信息。



7.2 目标文件

可以选择插入页分割行，并且定制风格行。

点击底部的“打开”按钮，程序将打开生成的文本文件。

7.3 预览

填写“预览页”、点击“预览”按钮，则程序提取预览页的文字后显示此文件。

```

1 new 1 x new 2 x snapshots_nav.html x README.md x snapshots_pdftools_extract_image.html x Your First Cup An Introduction to the Java EE Platform.txt.txt x
108 4.5 Building, Packaging, Deploying, and Running the firstcup-war Web Application ..... 4-11
109 4.5.1 Build, Package, and Deploy the firstcup-war Web Application..... 4-11
110 4.5.2 Run the firstcup-war Application .....
111 5 Next Steps
112 5.1 The Java EE Tutorial .....
113 5.2 More Information on the Java EE Platform.....
114 -----3 / 32-----
115 Preface
116 This is Your First Cup: An Introduction to Java Platform, Enterprise Edition, a short tutorial
117 for beginning Java EE programmers. This tutorial is designed to give you a hands-on
118 lesson on developing an enterprise application from initial coding to deployment.
119 Audience
120 This tutorial is intended for novice Java EE developers. You should be familiar with
121 the Java programming language, particularly the features introduced in Java Platform,
122 Standard Edition 7. While familiarity with enterprise development and Java EE
123 technologies is helpful, this tutorial assumes you are new to developing Java EE

```

8 提取 PDF 中的文字-批量

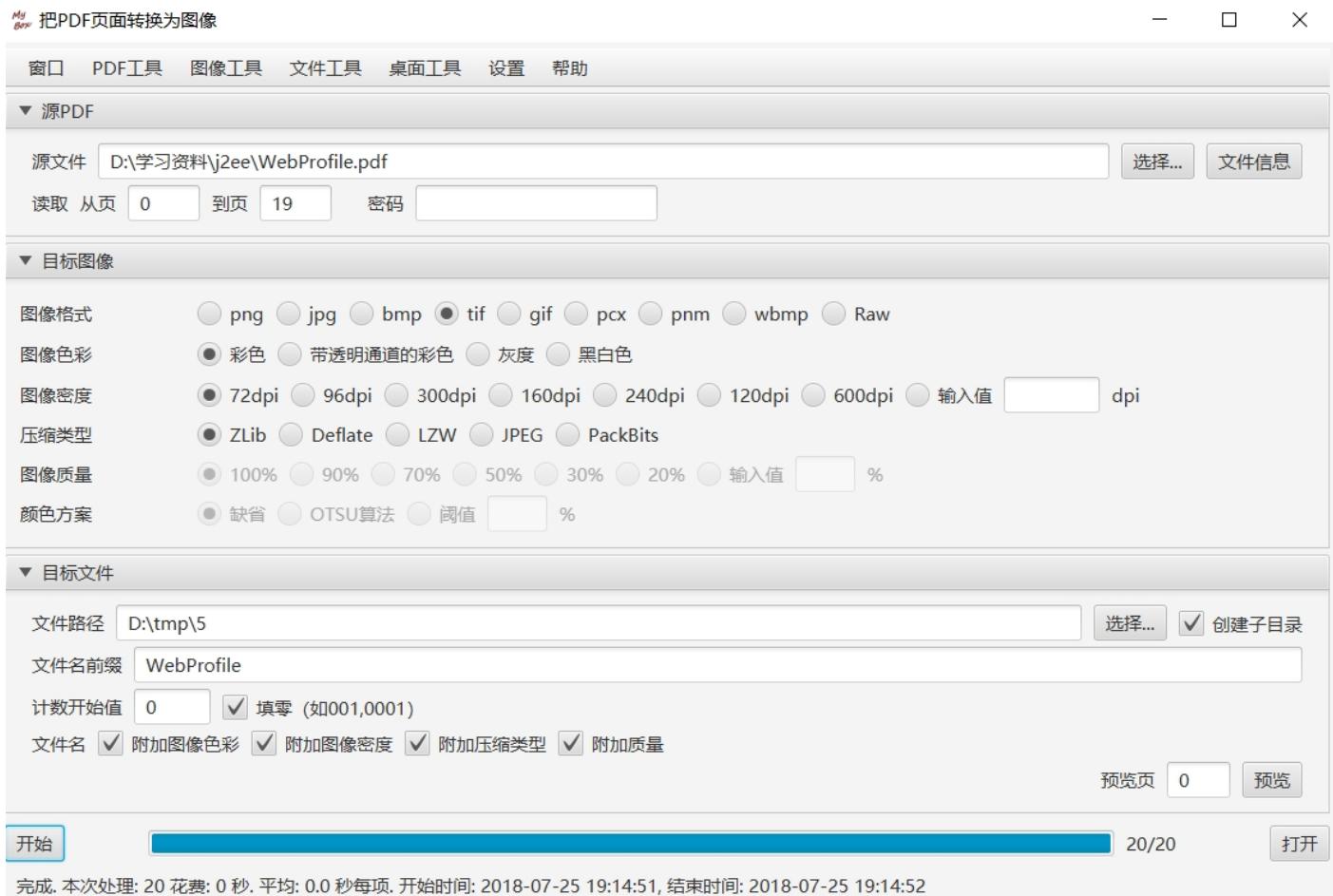
可以对多个 PDF 文件执行批量提取文字：



9 将 PDF 页面转换为图像

9.1 源 PDF

可以填写要处理的页码范围，若打开 PDF 需要密码则应当填写密码。



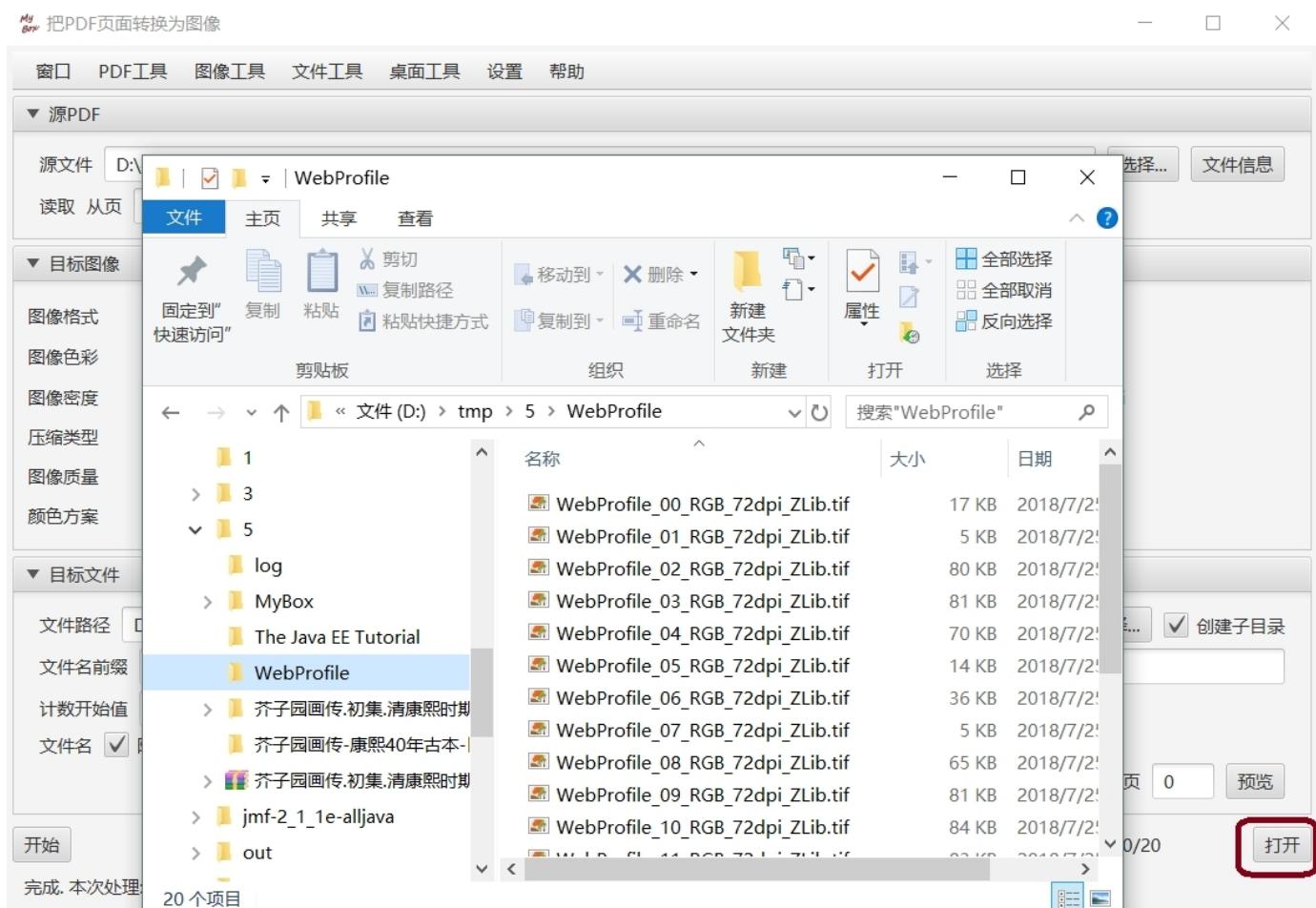
9.2 目标文件

若选择“创建子目录”，则程序将在目标目录中以 PDF 文件名创建一个子目录，将转换出来的图片放入此子目录。若不选此项，则转换出的图片直接放在目标目录下。

转换的图片文件名是：用户填写的“文件名前缀”（被自动初始化为 PDF 文件名）+ 计数值 + 图片扩展名。用户可以设置计数开始值。

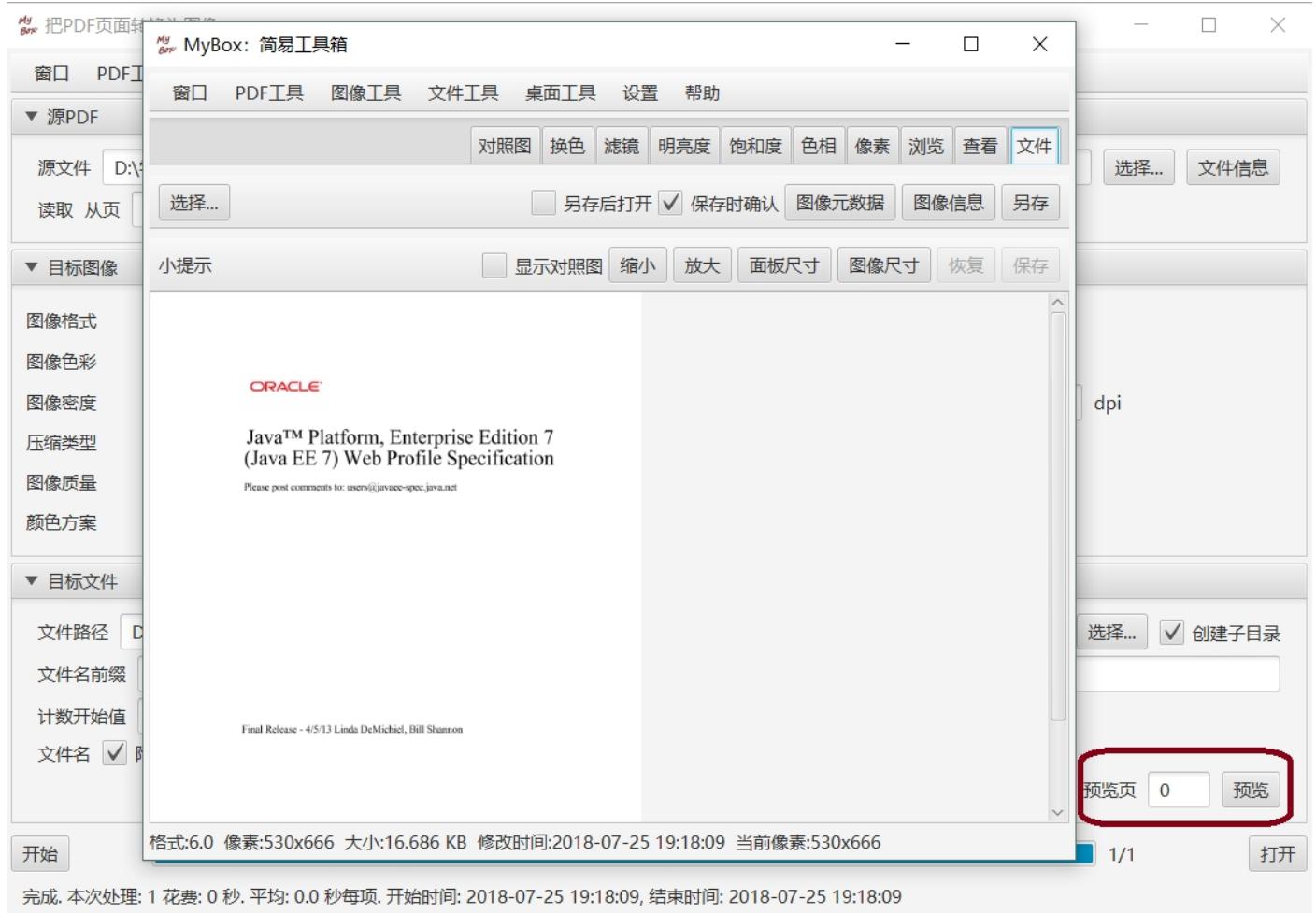
用户可以选择文件名是否附加图像色彩、图像密度、压缩类型、质量信息。

点击底部的“打开”按钮，可以查看图片在目录文件的生成情况：



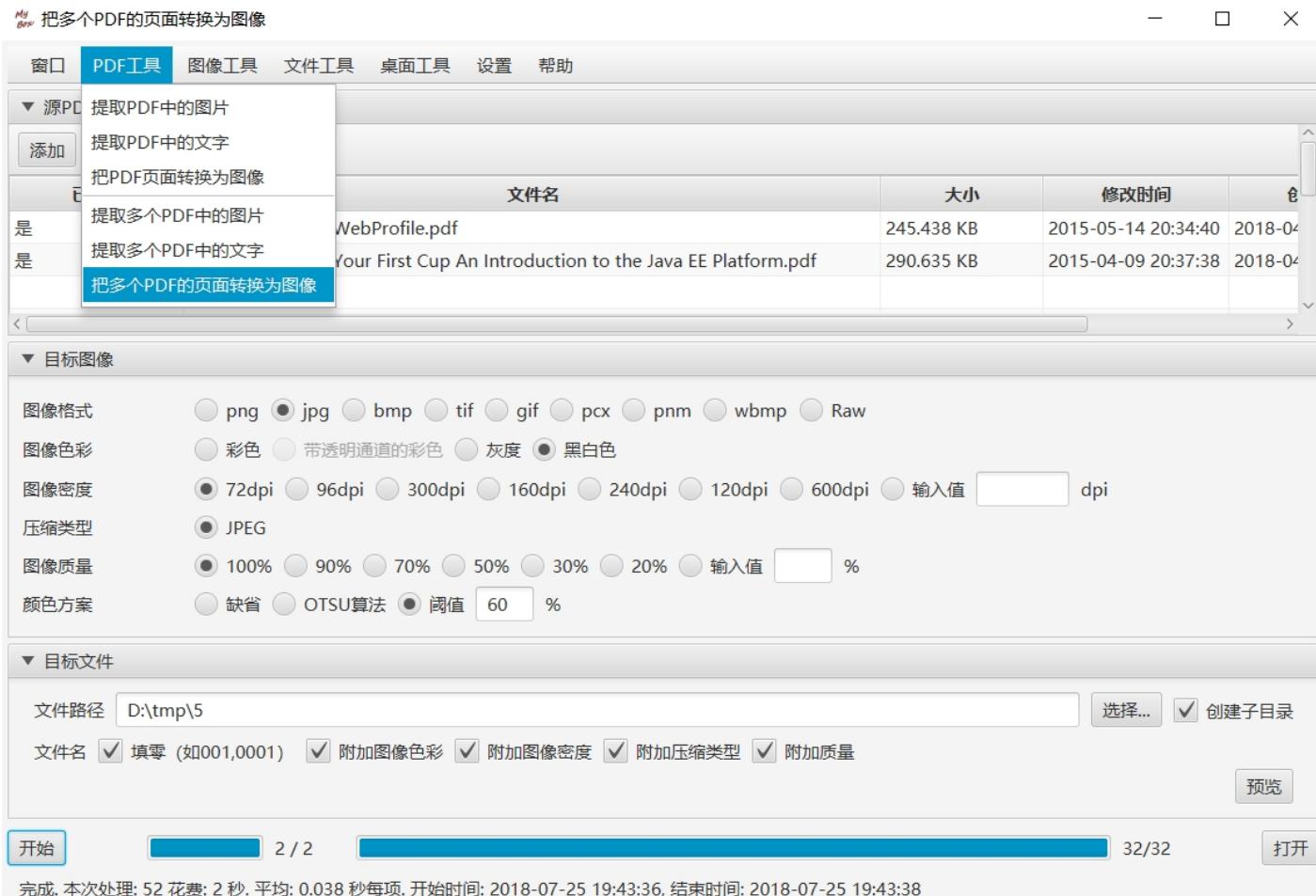
9.3 预览

填写“预览页”、点击“预览”按钮，则程序转换预览页后显示此图片。



10 将 PDF 转换为图片-批量

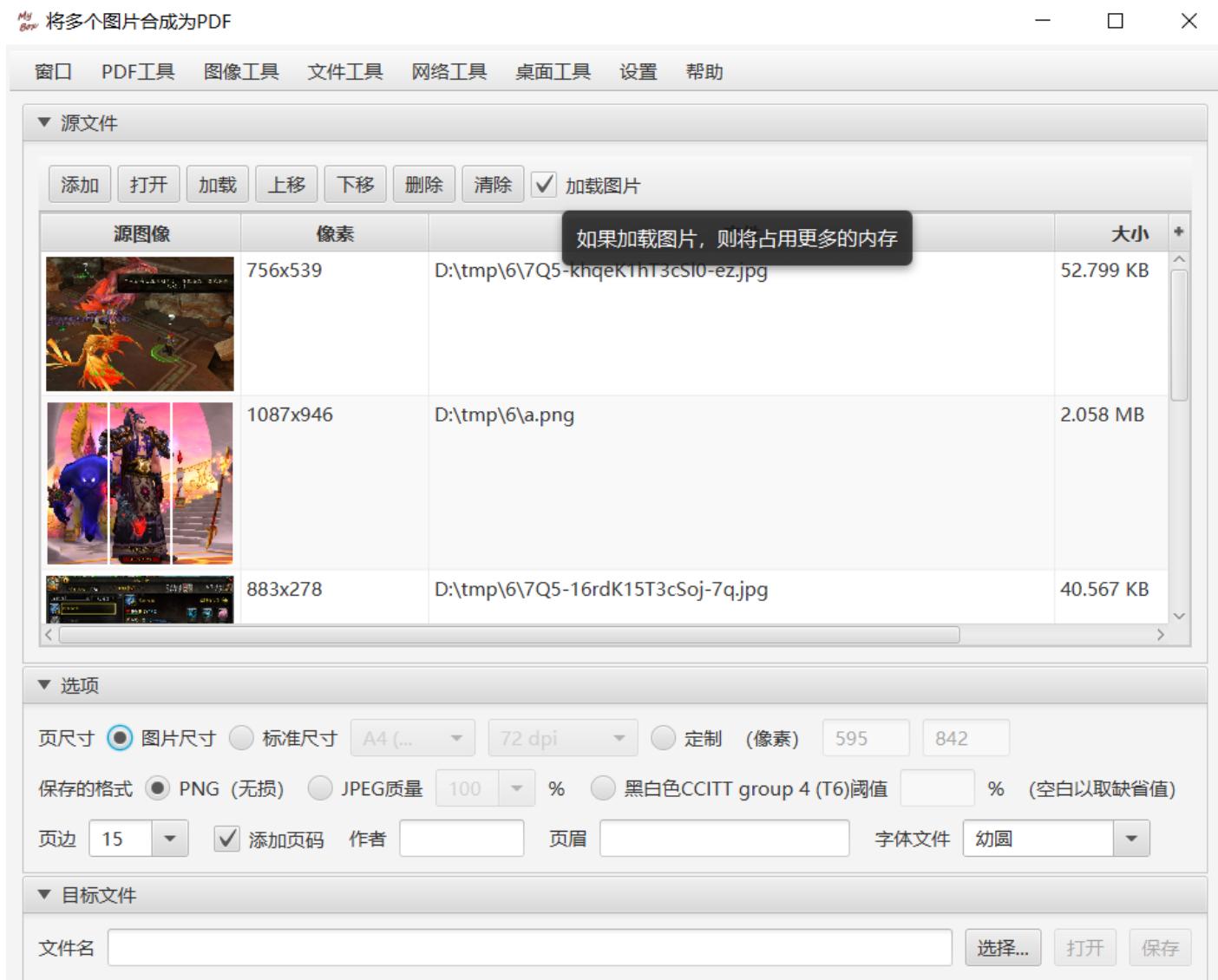
可以将多个 PDF 文件批量转换为图片：



11 将多个图片合成 PDF

11.1 管理源图片的列表

- 1) 添加、删除、清除、排序源图片。
- 2) 打开、加载、上移、下移所选择的图片。
- 3) 选择在添加图片时是否加载图片。加载图片会占用内存。

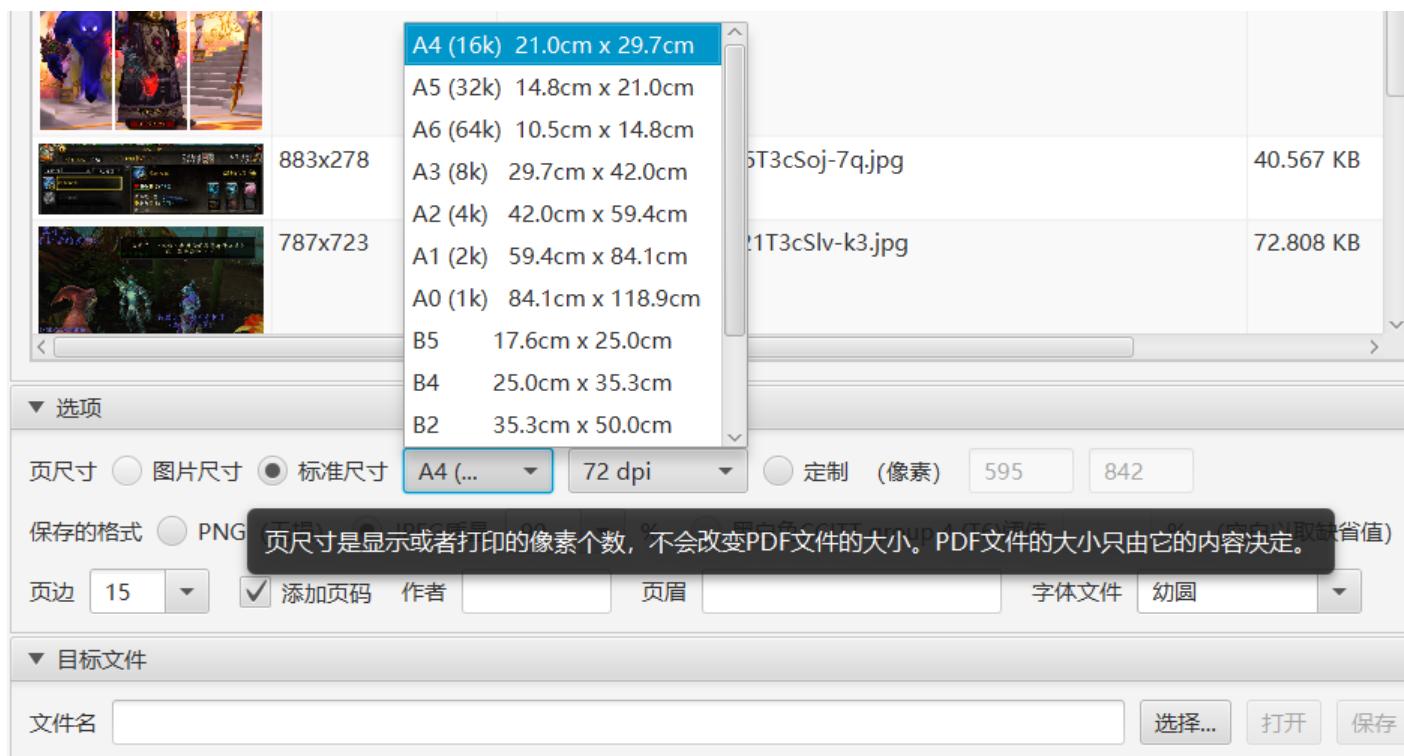


11.2 设置 PDF 的页面尺寸

可以设置生成的 PDF 的页面尺寸：

- 1) 以图片尺寸为页面尺寸。这样每页的尺寸都可能不同。
- 2) 选择标准纸张尺寸和图像密度，来决定页面的像素值。
- 3) 定制页面的尺寸（输入长宽的像素值）。

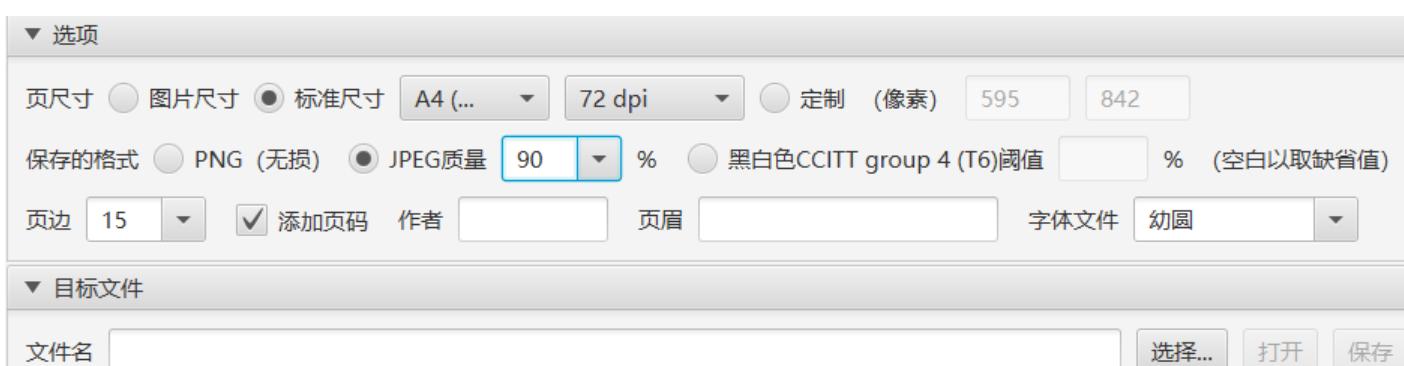
注意：PDF 页面尺寸用于显示或打印，对文件大小没有影响。PDF 文件大小只与它的内容有关。



11.3 设置图片压缩格式

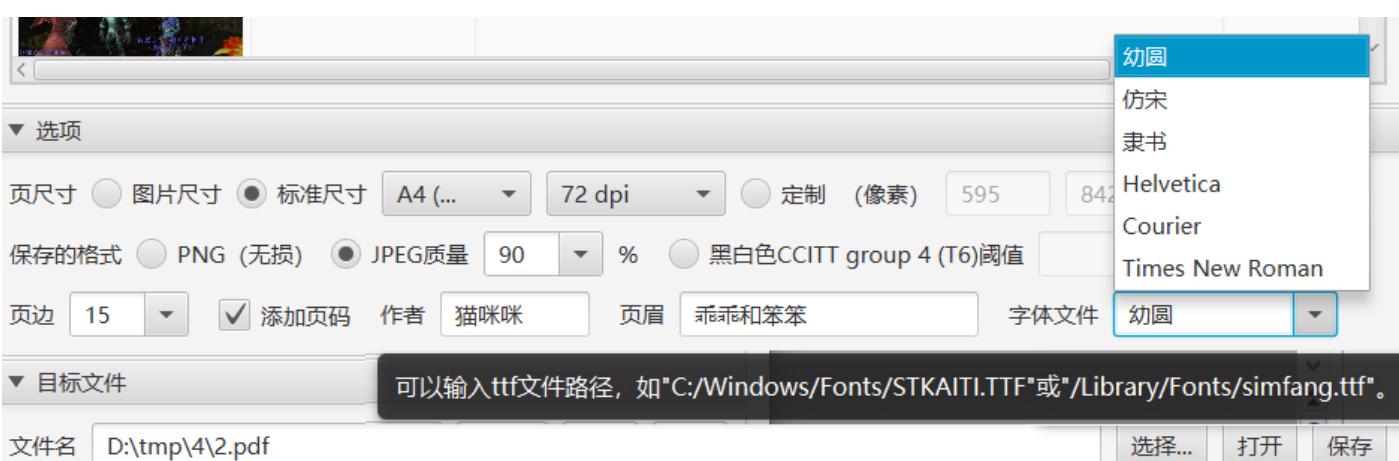
可以设置图片保存在 PDF 文件的格式：

- 1) 无损的 PNG 格式。
- 2) JPEG 压缩，可设置图片质量。
- 3) 以 CCITT group4(T6)压缩算法转换为黑白色，可设置转换阈值。

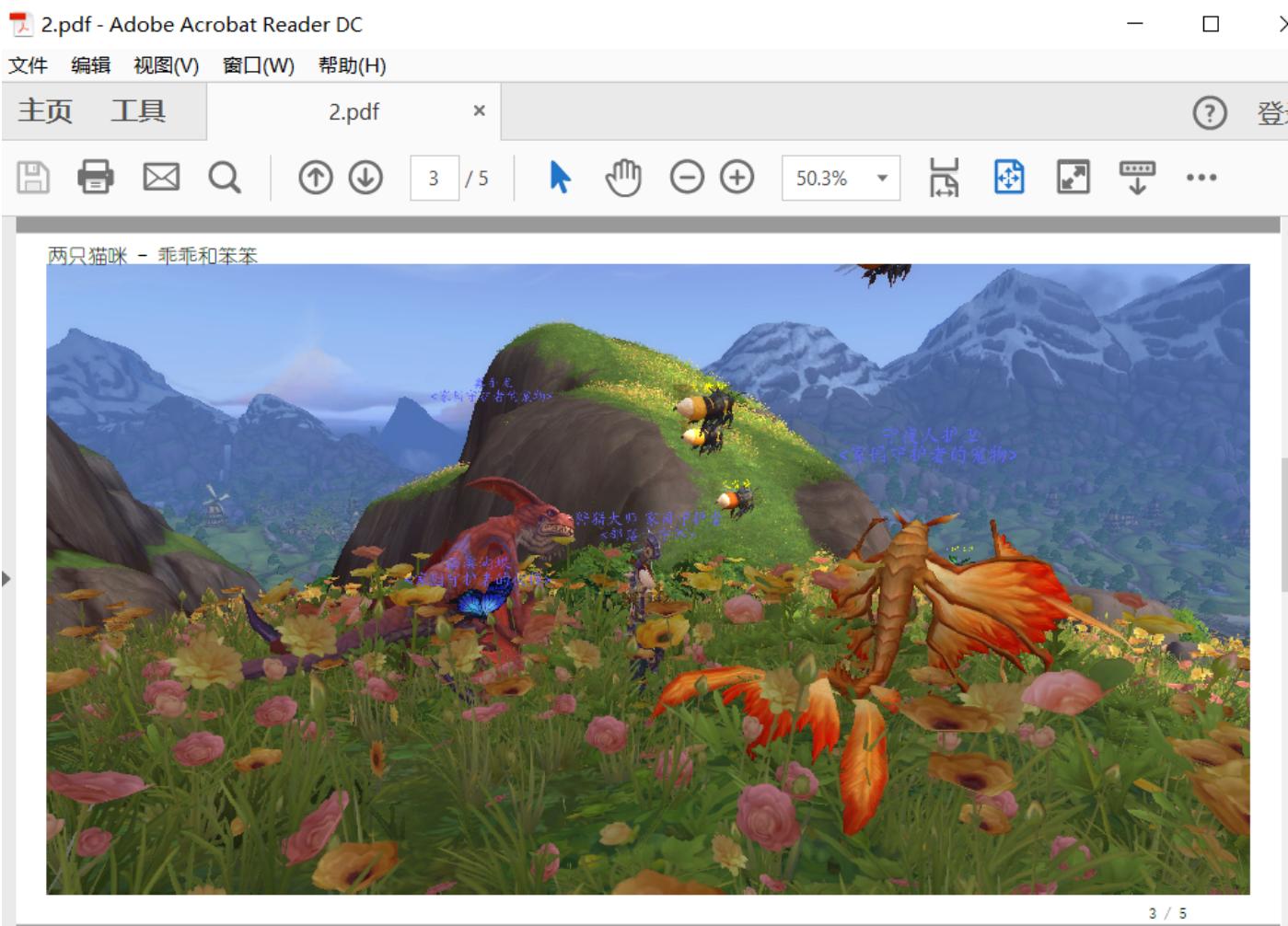


11.4 设置页边、作者、页眉

- 1) 设置字体以正确写入中文。程序自动查找系统中已安装的字体文件。
- 2) 若缺省字体选项不满足要求、或者字体在系统中定位失败，用户可以输入 ttf 字体文件的路径，如"C:/Windows/Fonts/STKAITI.TTF"或"/Library/Fonts/simfang.ttf"。



以下是合成多个图片生成 PDF 的一个例子：



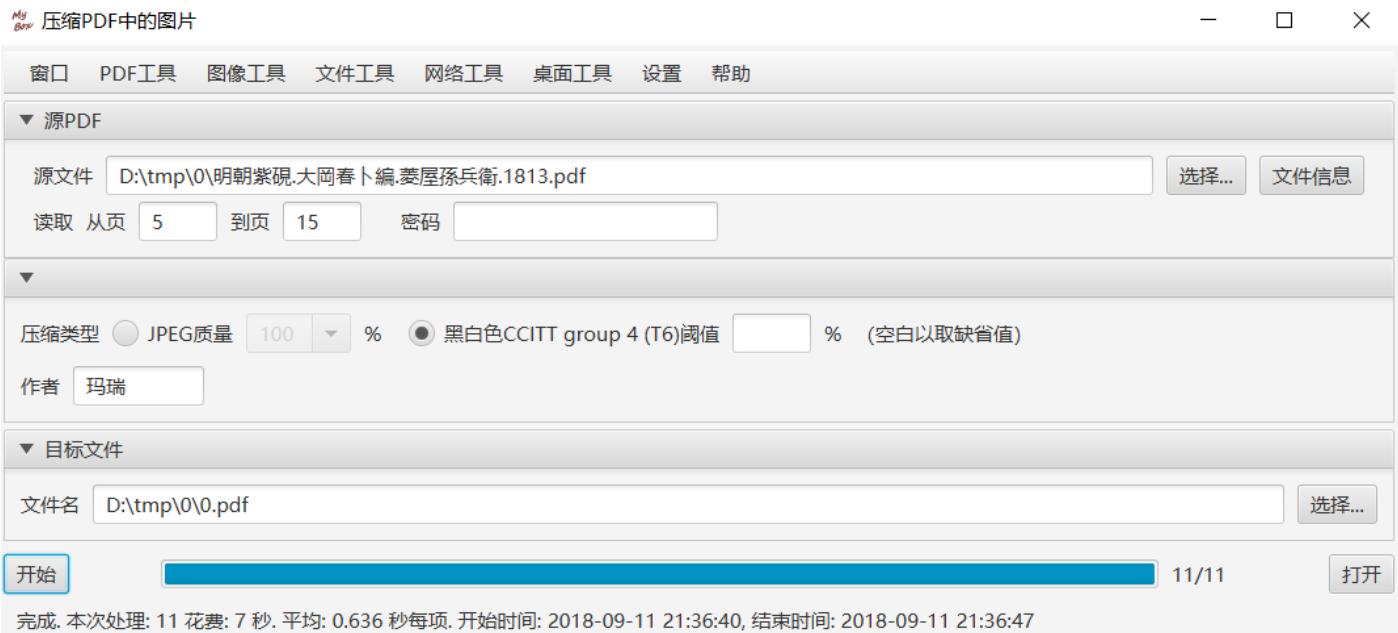
12 压缩 PDF 的图片

1) 设置要处理的起止页码。

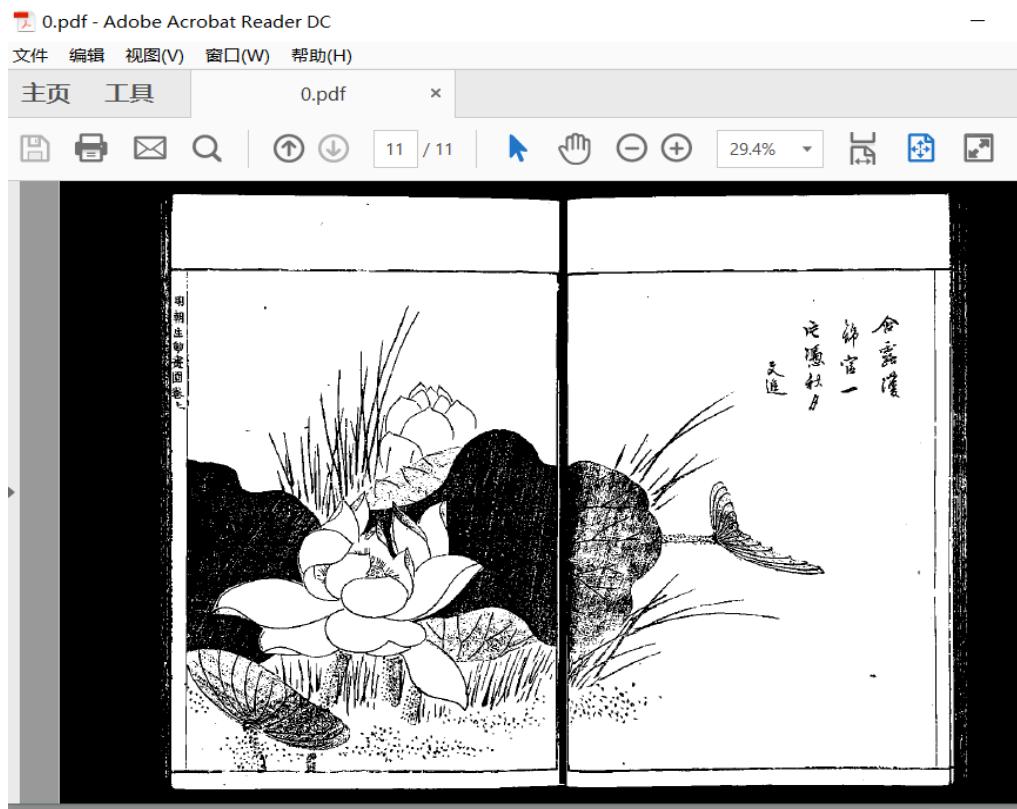
2) 设置压缩类型：

(1) JPEG 压缩，可以说设置图片质量。

(2) 以 CCITT group4(T6) 实现的黑白色压缩，可以设置转换阈值（空白以取缺省值）。



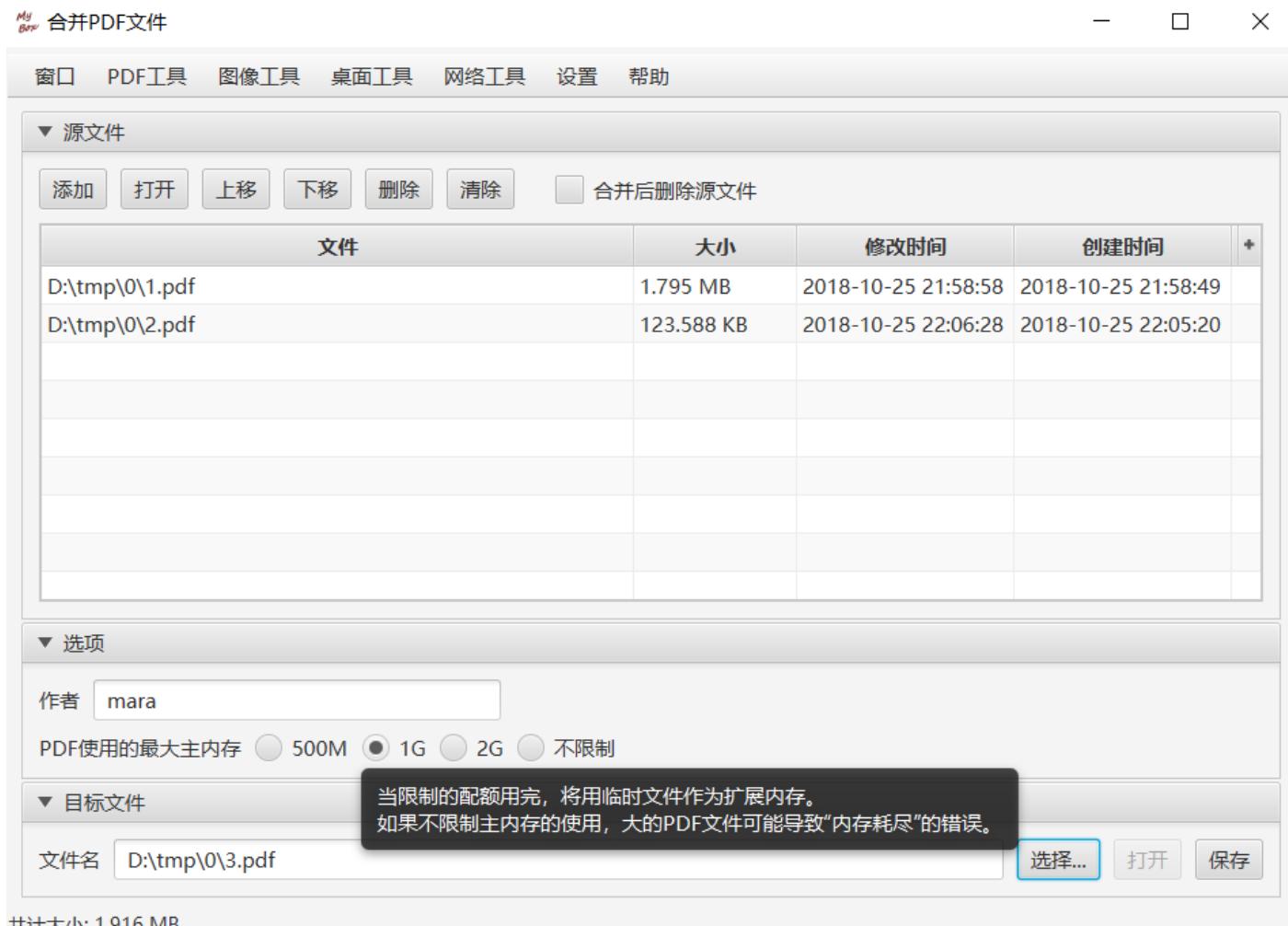
3) 源 PDF 中所选页面中所有图片将被转换为指定格式，然后保存在新的 PDF 文件中。



13 合并 PDF

- 1) 添加、删除、清除、排序源 PDF 文件。
- 2) 打开、上移、下移所选择的 PDF 文件。
- 3) 可以填写“作者”。

为了避免 PDF 文件太大造成内存溢出，用户需要注意设置“PDF 处理的最大主内存”和“临时文件目录”。



14 分割 PDF

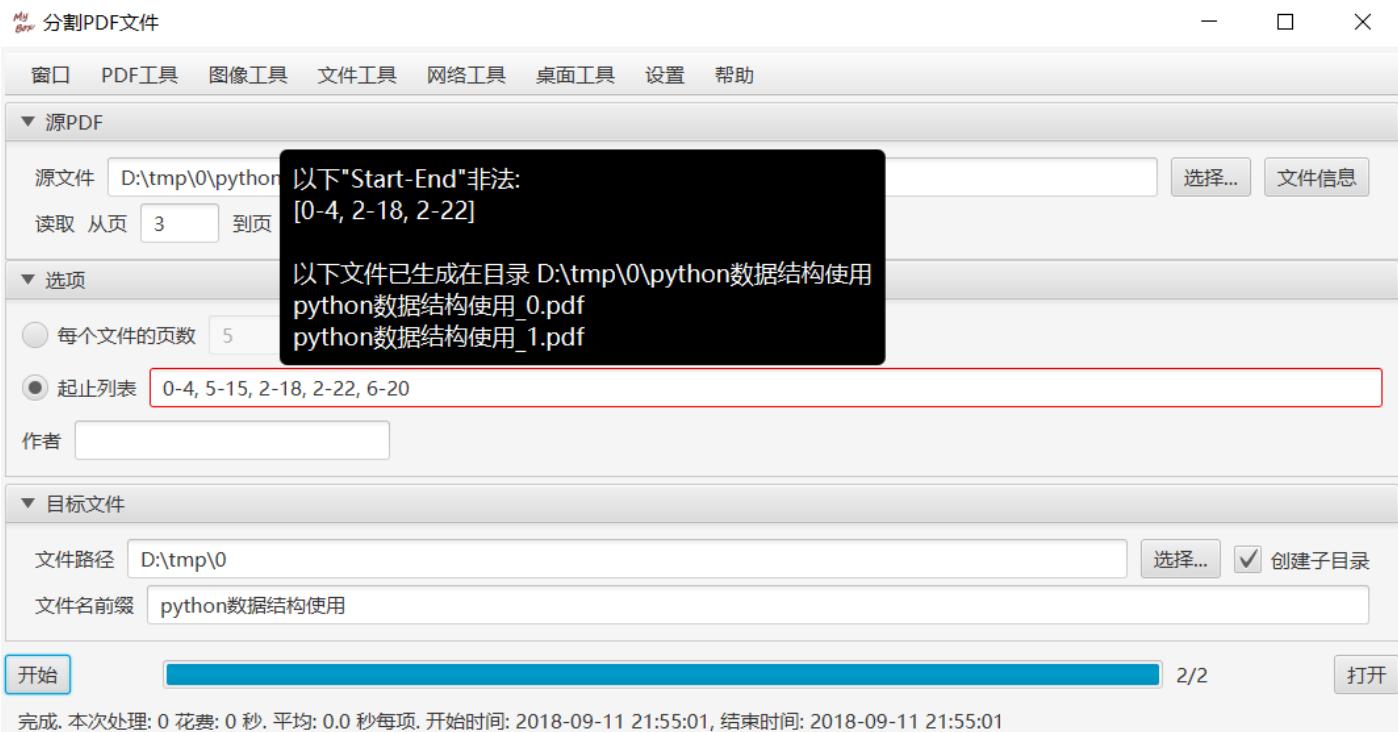
- 1) 设置要处理的总的起止页码。
- 2) 选择分割方式：
 - (1) 按页数均分，设置每个文件的页数。
 - (2) 按文件数均分，设置文件数。

(3) 按起止页码列表分割，格式为：以逗号分隔的“开始页码-结束页码”的列表。列表项不必按顺序排列，项与逗号之间可以有空格。若分隔项不在总的起止页码范围内、或者跨越了总起止范围，则被认为是非法的。非法的分割项不影响合法的分割项的执行。



- 3) 输入文件名前缀，程序自动在文件名后添加文件序数。

MyBox 用户手册 - PDF 工具 v5.0



<文档结束>