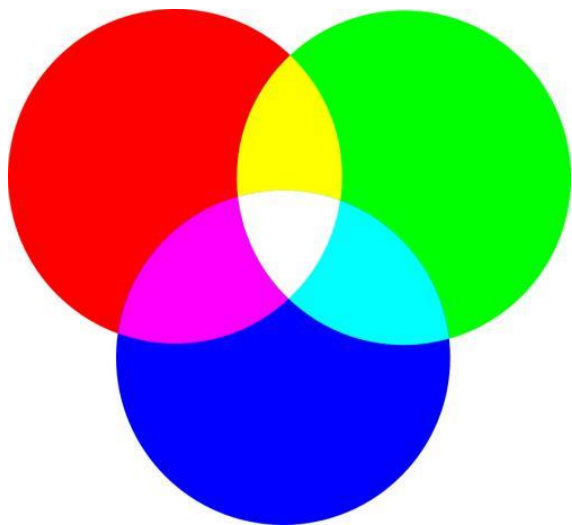


# <VideoSubfinder教程>

## 明确概念

### RGB

- 我们所熟知的红蓝绿三原色，取值均在0-255



### LAB

- 一个理论上包括了人眼可以看见的所有色彩的色彩模式，弥补了RGB的不足

$L \in (0, 100)$

$a \in (-128, 127)$

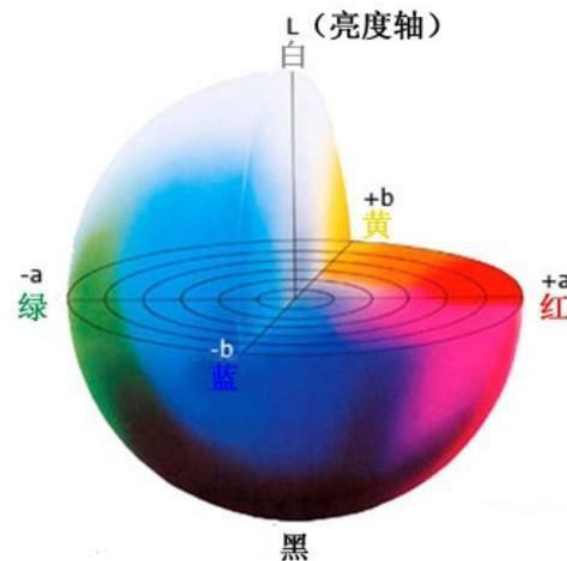
$b \in (-128, 127)$

为了便于计数，使其也同样处于0-255范围

$L = L * 2.55$

$a = a + 128$

$b = b + 128$



## 明确概念

### RBG图片

- 可以称为全彩色图片。
- 从一部彩色电影中截取一张图片，就是RGB图片



### 二值化图片

- 二值化，是使整个图像呈现出明显的黑白效果的过程。
- 截取图像中字幕所在的区域，二值化后就只剩下了黑色的字幕部分



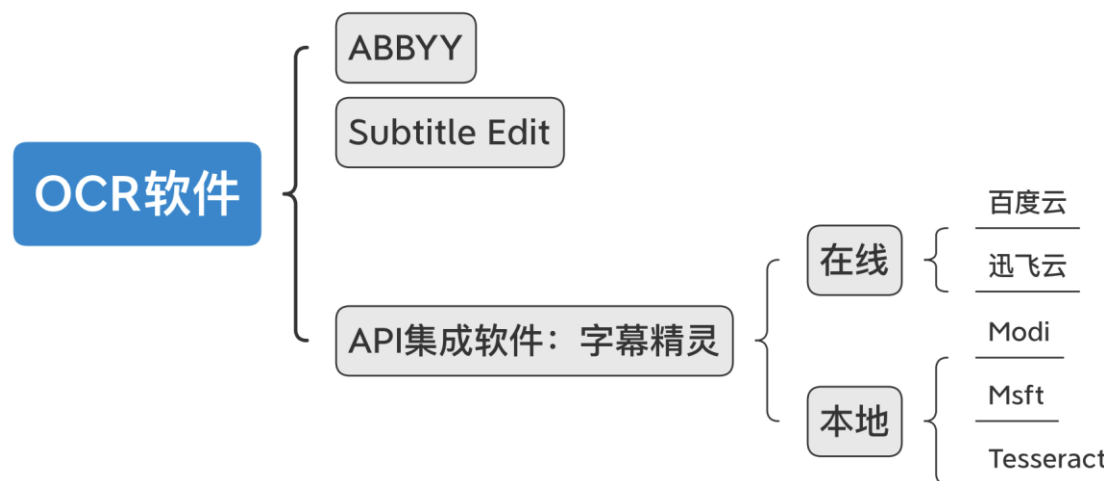
请记住我

# 明确概念

## OCR(光学字符识别)

- 简单来说就是把图片上的文字转为可编辑文本的过程
- 之所以对图片二值化处理，就是为了使OCR识别的结果更加准确

## OCR软件



### 注意事项:

- 使用在线API接口的OCR软件，可以直接识别RGB图片，不需要将图片二值化
- OCR软件的使用请查看相应软件对应的教程

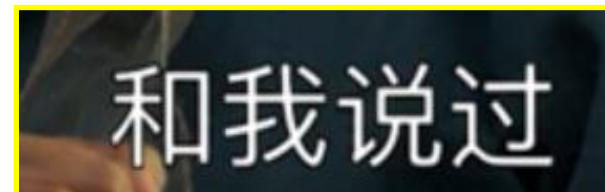
# 明确概念

## 字幕颜色

- 字幕颜色分为字体颜色和边线颜色
- 图中白色的为字体，紫色的为边线
- 但在日常使用中，字体颜色由于常用，字幕颜色通常就指字体颜色



- 大部分视频字幕都没有边线或者边线不明显，比如国产剧中的字幕



## 关于程序

程序主要提供两个功能：

- 1) 通过数字图像处理算法，自动检测视频中带有硬字幕的帧，生成只带有时间轴的空白字幕；
- 2) 通过文本挖掘算法，用带有硬字幕的图片生成图形文字。可以用其他软件来OCR，生成既有文本又有时间轴的字幕。

### 注意事项：

- 为了使程序正常运行，需要下载安装"Microsoft Visual C++ Redistributable for Visual Studio 2015, 2017 and 2019",  
下载地址：  
<https://support.microsoft.com/en-us/help/2977003/the-latest-supported-visual-c-downloads>  
x64: [https://aka.ms/vs/16/release/vc\\_redist.x64.exe](https://aka.ms/vs/16/release/vc_redist.x64.exe)  
x86: [https://aka.ms/vs/16/release/vc\\_redist.x86.exe](https://aka.ms/vs/16/release/vc_redist.x86.exe)  
最新版本测试基于Windows 10。
- 软件不支持阿拉伯语以及手写字体

## 快速入门指南

### 第一步：通过视频截图，生成RGB图片

- 1) 点击菜单栏中"File->Open Video"（推荐使用OpenCV，提取的时间轴准确性更高）；
- 2) 在"Video Box"中，通过移动垂直和水平分隔线来缩小检测区域，结果更准确、时间轴更完整；



### 注意事项：

- 移动垂直分隔线，需保证左右移动的距离相同，不然后续识别很可能会漏掉部分时间轴



# VideoSubfinder教程

## 快速入门指南

- 3) 确认字幕在检测区域内的对齐方式: "Center/Left/Right/Any" 中心对齐/左对齐/右对齐/任意对齐, 在"Settings"标签页里的"Text Alignment"中选择对应属性



### 注意事项:

- 尽量使字幕位于检测区域的水平中心
- 任意对齐的稳定性不如其他几种对齐方式



## 快速入门指南

- 4) 强烈推荐使用"Use Filter Colors"颜色滤镜提高识别结果准确率，该步骤可以跳过：

\* - 拖动进度条，找到带有字幕的图像

\* - 鼠标点击"Video Box"任意位置，选中后按"U"，全屏查看图片，鼠标左键点击字幕像素点，即可获取字幕的颜色信息

\* - 字幕颜色信息在"Settings"便签页中的右下角，复制Lab颜色信息到左边的"Use Filter Colors"编辑栏中，如：Lab: l:0 a:128 b:128

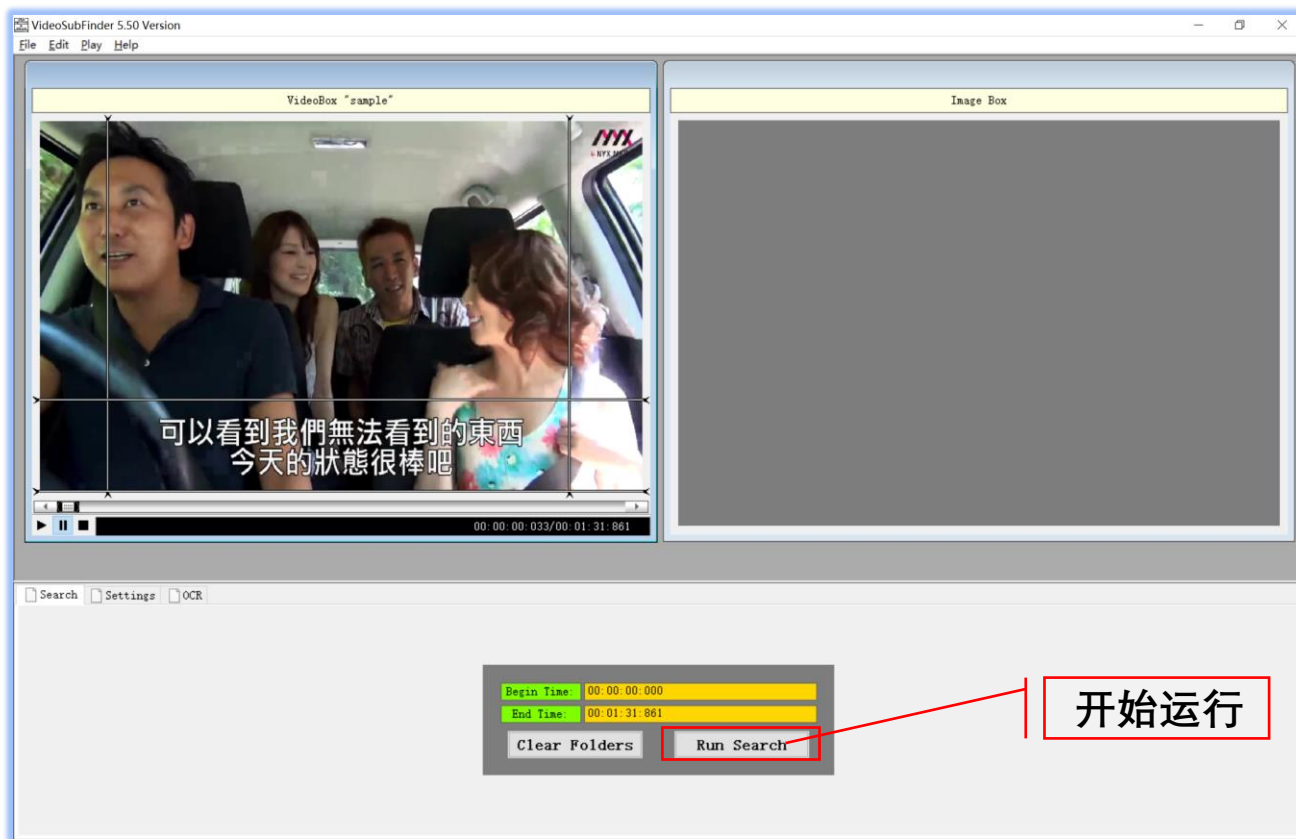
\* - 如果字幕颜色有多种，可以用“Ctrl+Enter”在“Use Filter Colors”编辑栏中添加新的一行颜色信息，并打开"Combine To Single Cluster"选项



# VideoSubfinder教程

## 快速入门指南

- 5) 在“Search”标签页中点击“Run Search”，即可生成RGB图片



### 注意事项:

- 可以在软件根目录的RGBImages文件夹查看截取的RGB图片
- 如果你只需要时间轴，这步之后，直接跳到“OCR”标签页，按“Create Empty Sub From RGBImages”，生成一个只有时间轴信息的字幕
- 如果你是用的使用在线OCR接口的软件，可以直接OCR，然后跳到第8点

## 快速入门指南

### 第二步：二值化RGB图片

- 6) 如果不使用颜色滤镜，则需要注意：

\* - 在继续下一步之前，检查字幕边线的颜色是否比字幕文本的颜色更深（大多数情况下是这样的，如果不是，取消选择在“Settings”标签页中的右边栏中的第一个复选框“Characters Border Is Darker”字幕边线颜色更深）

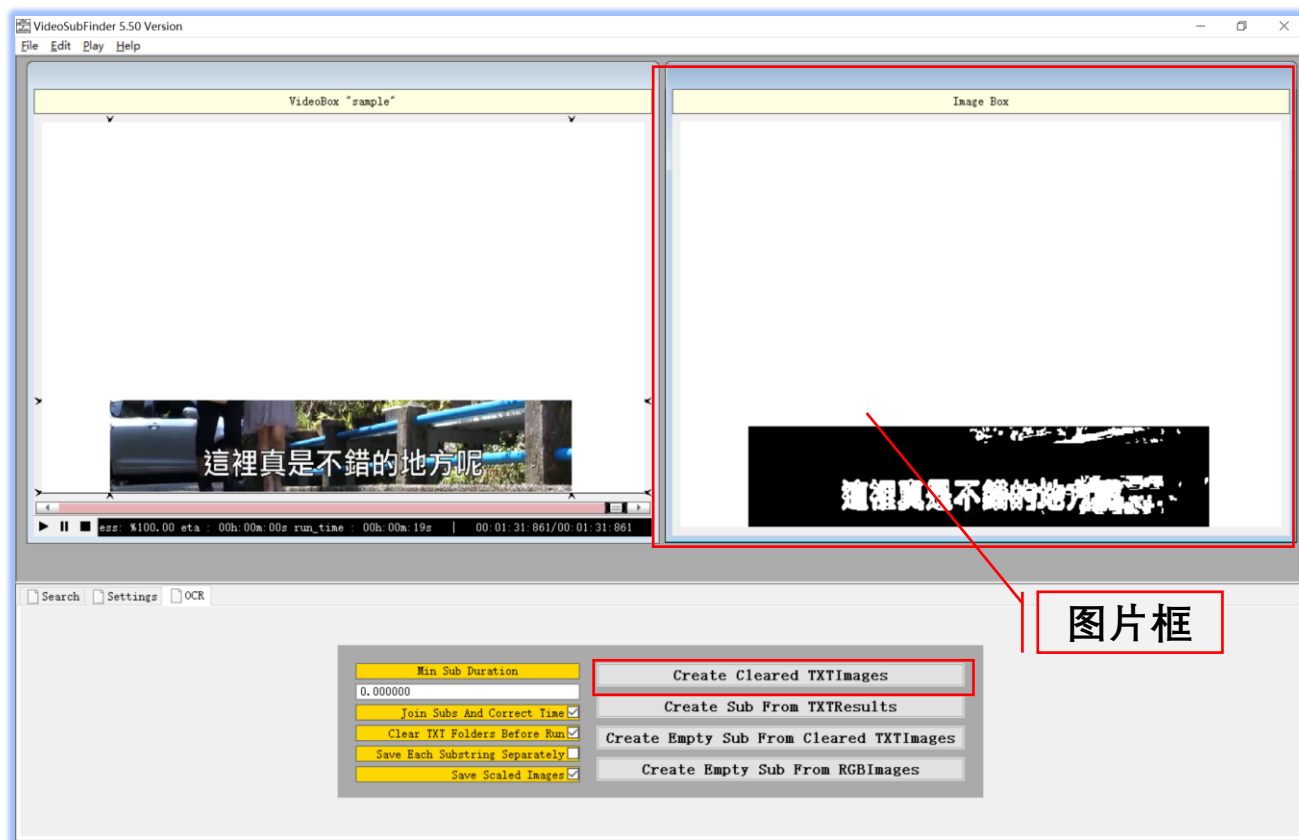
\* - 在大多数情况下，程序能正确地识别与字幕文本相关联的颜色，但某些情况过于复杂，就需要通过其他参数设置来处理，详见后面参数设置；



# VideoSubfinder教程

## 快速入门指南

- 7) 点击“OCR”标签页中的“Create Cleared TXTImages”，即可生成二值化图片



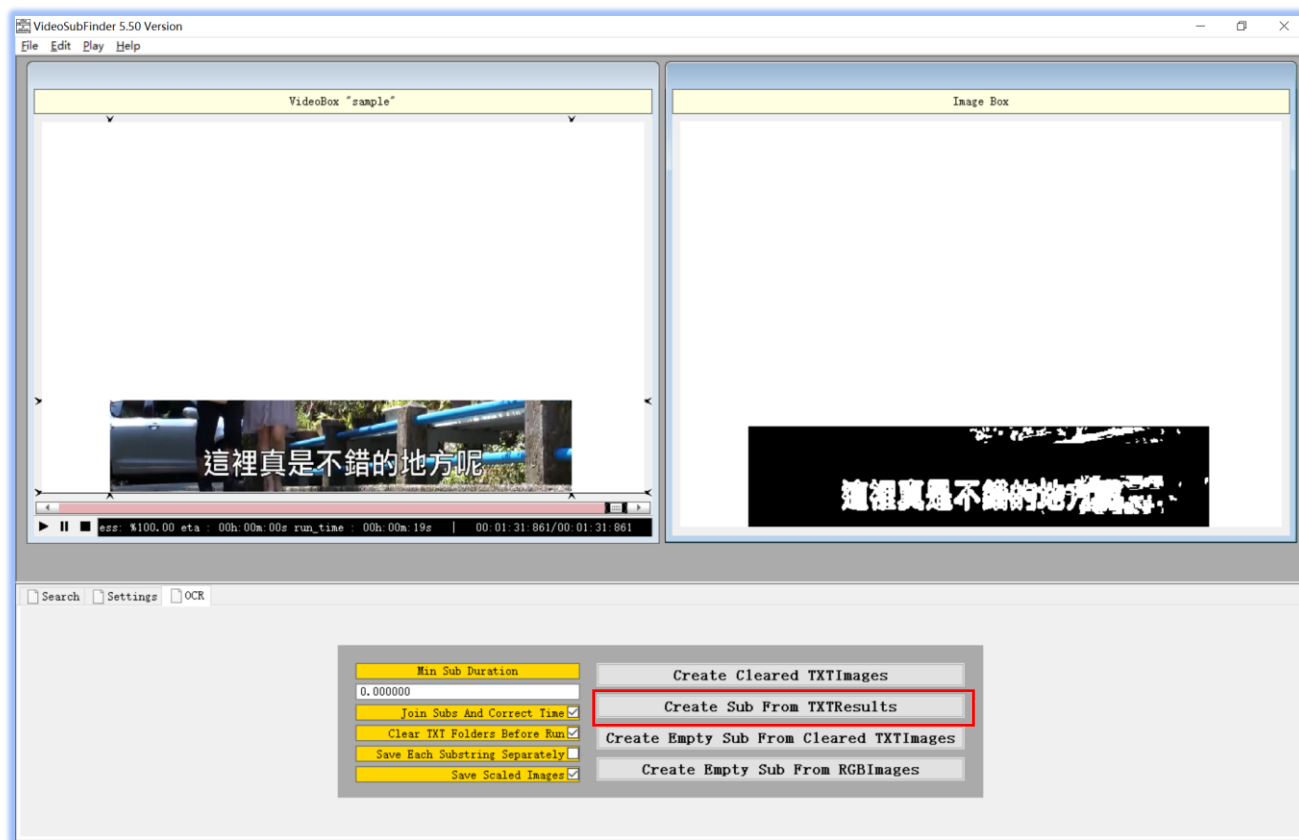
### 注意事项:

- 左边的“Image Box”图片框仅是预览，不是最后的二值化结果，最终的结果可以在软件根目录的TXTImages文件夹查看生成的二值化图片

# VideoSubfinder教程

## 快速入门指南

- 8) 通过OCR软件，将二值化的图片生成文本后，点击Create Sub From TXTResults，即可生成完整的字幕



### 注意事项:

- OCR的结果可以在软件根目录的TXTResults文件夹中找到

## 参数设置详解

<在"Run Search"生成RGB图片之前>

\*- 为了减少在"Run Search"生成RGB图片时出现的：丢轴、错轴；

在"Clean TXT Images"时出现的：文字未被识别、只识别部分的情况。

对"Moderate Threshold"参数在 [0.25, 0.6]范围内进行调整：

为了找到最佳的参数值，尽量多的选择视频中背景明暗度差异较大的帧来进行测试，

比如背景太亮或者太暗，在"Test"时，需要对'After First/Second/Third Filtration'不同的阶段进行测试，

直到在右边的"Image Box"图片框中出现完整的白色字幕。

0.25 - 大多数视频通用，特别是1080p的视频，但是会生成大量无用的图片

0.5-0.6 - 适用于字幕边线明显的，像白字黑边框的字幕，或者画质低于480p的视频

0.1 - 适用于字幕没有边线或者字幕颜色比较浅的视频，但是会生成大量无用的图片



## 参数设置详解

<在"Run Search"生成RGB图片之前>

\*- 时间轴的切分也受以下两个参数影响：

`vedges_points_line_error = 0.3`

`ila_points_line_error = 0.3`

0.3意为两个相近字幕的允许差异值为30%，比如：两条不同的字幕都只有十个字，在允许差异值为30%的情况下，其中一条字幕只要有七个字以上与另一条字幕相同，就视为同一条字幕。同理，如果允许差异值为50%，则只需要五个字以上与另一条字幕相同就可视为同一条字幕

值越高，时间轴分割的越少

值越低，时间轴分割的越多

### 注意事项：

- 有的参数会生成大量无用的图片，但是注意一个原则：宁愿多、不愿少，防止出现漏轴的情况。



## 参数设置详解

<在" Clean TXT Images"二值化RGB图片之前>

\*- 检查字幕颜色是否更深，如果不是，取消选择"Settings"标签页中右边第一项

"Characters Border Is Darker"的复选框

\*- 二值化过程中部分字符丢失

与"Moderate Threshold For Scaled Image"这个参数有关，参数范围[0.1-0.25]

0.25 - 大多数视频通用，如果有字符丢失，适当调低参数

0.1-0.15 - 适用于字幕边线不明显的或没有字幕边线的、字幕颜色与背景颜色相近的

推荐打开"Use ILAImages before clear TXT images from borders"选项，能改善结果

\*- 减少二值化过程中多余符号的产生

"Clear Images Logical"默认打开，但是可能会移除部分正确的字符

如果使用了颜色滤镜，以及提取的ILAImages质量很好，推荐打开"Use ILAImages for getting TXT symbols areas"选项，可以减少多余符号的生成

打开"Remove too wide symbols"选项，减少多余字符的产生

## 参数设置详解

### <颜色滤镜>

\*- 常用颜色滤镜参数范围

Lab: l:180-255 a:108-148 b:108-148 (大多数视频通用)

Lab: l:200-255 a:118-138 b:118-138

Lab: l:220-255 a:118-138 b:118-138 (强力颜色滤镜，视频质量很好的情况下使用，字幕亮度不稳定不推荐使用)

参数格式:

Lab: l:l\_val a:a\_val b:b\_lab\_val

Lab: l:min\_l\_val-max\_l\_val a:min\_a\_val-max\_a\_val b:min\_b\_lab\_val-max\_b\_lab\_val (亮度范围)

\*- 边线的颜色参数范围同上，除非边线很明显，否则"Use Outline Filter Colors"选项不建议使用强力滤镜

\*- 在“Video Box”中，按“T”可全屏查看设置颜色滤镜后效果，可通过左右方向键和空格键来查看不同帧

\*- 字体颜色是红色的、边线颜色是绿色的、字体和边线重合部分是黄色的

同样:

按"U"可全屏查看原始视频帧

按"Y"查看原视频中剥离出来的字体的颜色，比如：字体是黄色的，屏幕上应该就只出现黄色的字幕

按"I"查看原视频中剥离出来的边线的颜色

## 参数设置详解

### <二值化质量不高的视频>

#### \*- 字幕亮度不稳定的情况

打开"Extend By Grey Color", 人工设定"Allow Min Luminance", 最低亮度的值  
如果使用颜色滤镜, 建议把最低亮度的值设成跟颜色滤镜的"min\_l\_val"相等  
参数范围: [min("允许的最低亮度", 自动检测的最低亮度), 自动检测的最高亮度]

#### \*- 借助第三方图片增强软件

使用"Topaz Gigapixel AI": <https://topazlabs.com/gigapixel-ai/>, 对提取的RGB图片进行增强  
建议按默认的两倍放大增强, 由于图片的名称就是字幕的时间轴, 增强后的图片名称应与原来的一致

#### \*- "Video Gamma"和"Video Contrast"对结果也有帮助

某些情况下, 设置"Video Gamma" == 0.7、"Allow Min Luminance" == 100、打开"Extend By Grey Color", 会很有帮助  
具体情况中, 需要自己不断测试才能得到最优的参数值

## 参数设置详解

### <二值化没有边线的字幕>

\*- 这种情况下，某些字符的分离会变得很困难

可以打开“Use ILAImages before clear TXT images from borders”选项；

正确设置合适的字幕颜色滤镜；

并且，在背景是动态的情况下，ILAImages会很有帮助；

以下列举的参数对背景分离也很有帮助：

`min_sum_color_diff = 0`

`moderate_threshold = 0.1`

`moderate_threshold_for_scaled_image = 0.1`

`use_ILA_images_for_getting_txt_symbols_areas = 0`

### <提高软件运行时的性能>

\*- 默认的参数为：

`"CPU kmeans initial loop iterations" == 20`

`"CPU kmeans loop iterations" = 30`

以上参数适应于各种情况，特别是在字幕没有良好的边线的情况下，当字幕有明显的边线的时候，两个指标的参数都可以设为10。

VideoSubfinder教程

# 参数设置详解

<常用术语>

ISAIImages - Intersected Subtitles Areas 交叉字幕区域

ILAIImages - Intersected Luminance Areas 相交亮度区域

# <翻译：豆瓣@我在成都养熊猫>

*\*- 相较于英文版教程，进行了部分筛减和补充，完整版请查看英文版本*

*\*- 欢迎加入硬字幕提取交流群：1161766056*