FileManager使用手册

Version: 3.38.6

Github: https://github.com/codecyou/FileManager

前言

由于以前硬盘文件有很多重复文件,为了整理文件去除重复文件,使用过CCleaner里面的查找重复文件功能,但是并不好用,用着感觉挺费劲的,于是想着自己写一个程序用来查找重复文件。

为了方便计算文件hash值,又新增了"计算hash值"程序。

Windows 10 不知道哪个版本之后,搜索文件功能变得不太好用,有时候会自动切词、或者会自动正则匹配这种情况,有时候只是想搜一个确切的具体文件名,然后结果就会出现很多不相关的内容,又不想使用everything,于是新增了"搜索文件"程序。

后来为了在硬盘之间备份同步文件,于是自己又写了文件同步备份的程序(最初想法就像阵列 卡那样),可以自动比对两个目录下的文件差异,然后执行同步备份或者还原的操作。

为了方便校对项目不同代码版本区别,新增了"比对文本文件内容"程序。

再后来整理文件的时候需要用到重命名功能,之前使用"灵者更名"和"Bulk Rename Utility",都挺好用的,但是"灵者更名"不能直接递归操作子文件夹下的文件和目录,"Bulk Rename Utility"界面有点复杂,当时使用正则匹配的时候没弄明白,加上文件名不能按自然数排序,会出现1.jpg,11.jpg,2.jpg这种情况,于是自己写了批量重命名的程序。

再之后,曾经百度下载过很多照片、写真,发现几千张照片重复的,比如分辨率不一样、文件 大小不一样、相同图片只是裁切过的、相同图片但是调过色的、相同图片但是有些有水印的,这种 情况下就无法用之前的文件查重程序筛选了,于是又新增了"查找相似图片"的程序。

后来发现硬盘上有很多之前刷抖音随手保存的视频,很多视频也是重复的(有的是重复发布,有的是不同抖音账号互相搬运的,视频内容一致,但是抖音水印不一致,而且分辨率也大多不一样),这种情况也无法使用之前的"查找重复文件"程序查找筛选,于是又新增了"查找相似视频"的程序。

后来发现缓存的一些视频,比如早期抖音app缓存的视频,很多是不完整的,于是新增了"查找损坏/不完整的视频"程序。

为了去除很早之前保存的抖音视频中最后三秒的那段片尾,于是新增了"视频截取"程序,该程序也可以用于截取视频片段。又新增了"音频处理"程序,用于从视频提取音频、从音频截取音频片段以及转换音频格式。

为了方便修改文件和目录时间戳,新增了"时间戳操作"程序。

有很多图片是WEBP和HEIC格式的,查看不方便,于是新增"图片处理"程序,用于检测图片真实数据类型,并可以进行简便图片数据类型转换。

程序功能介绍

"查找重复文件":用于查找重复文件,比较的是文件名、文件大小、修改时间,支持按文件后缀名选中或者排除要查找的文件类型。

"查找重复文件(hash方式)":用于查找重复文件,比较的是文件hash值。

"文件备份与同步":用于同步与备份文件,校对两个目录下的文件差异,同步文件修改操作,或者用备份还原。支持时间偏移修正(不同文件系统保存时间戳精度不同)、支持通过校验hash值确定是否为相同文件。使用简单方便,删除文件操作默认启用安全删除(删除的文件会移动到文件所在根目录下的safe_del目录,类似windows的回收站),如果想要更加强大的功能体验,推荐使用"Beyond Compare"。

"还原文件": 用于根据new old record记录还原文件或者重新操作文件。

"清除空文件夹":用于递归清除指定目录下所有空文件夹。

"搜索文件或目录":用于搜索文件或目录,支持1.按文件名,2.按文件大小,3.按文件数据类型,4.按记录文件名或者路径的txt,5.按视频时长、音频时长,6.按视频、图片分辨率,7.按文本内容,搜索文件或目录,支持正则匹配、条件匹配,支持按时间过滤,支持对搜索到的结果进行过滤(正则匹配),可以直接将搜索到的项目移动到指定目录(支持将所有项目移动到一个目录下或者保持原来的目录结构),支持根据文件真实数据类型搜索文件。使用简单方便,如果想要更加强大的功能体验,推荐使用"Everything"。

"拷贝目录结构":用于将一个目录的目录结构导出到记录文件,或者直接拷贝到另一个目录。(注意:只会拷贝目录结构,不会拷贝文件!)

"比对文本文件内容":用于比对两个目录下的同名文件的内容,或者比对两个文本文件的内容。这个功能原来是为了方便比对不同代码版本的代码差异而写的。当输入的是两个目录时只会比对两个目录下的同名文件,当输入的是两个文件时,则不会关注两个文件是否同名,会直接比对两个文件的内容差异。使用简单方便,如果想要更加强大的功能体验,推荐使用"Pycharm"自带的代码比较功能。

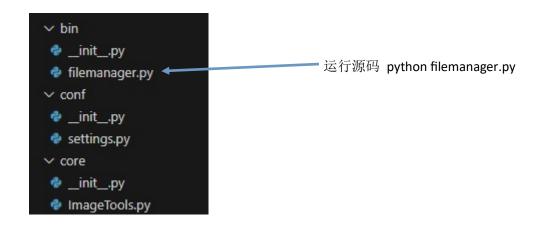
"计算hash值":用于计算文件的hash值。支持批量操作(拖拽多个文件到程序界面,会计算所有选中文件的hash值)、递归操作(拖拽文件夹到程序界面,会计算该文件夹下所有文件的hash值),并且如果是可执行程序会获取该程序的版本号。还可以从计算结果中查找输入的字符(用于验证hash值)。

- "提取视频帧图像":用于从视频中提取某一帧的图像或者某一秒的图像。
- "查找相似图片":用于查找相似的图片。
- **"查找相似视频"**:用于查找相似的视频,原理:比较两个视频某一秒的图像是否相似,也就要求两个视频可以分辨率不一样,但是时间轴不能有变化(至少是在程序比对的那一秒的时间轴应该是一致的)。
 - "以图搜图":用于从一个目录中搜索指定的图片。
 - "以视频搜相似视频":用于根据视频在指定目录中搜索出与该视频相似的视频。
- "批量重命名":用于执行批量重命名。支持普通搜索、精确搜索、正则匹配三种模式,支持替换字符、插入字符、插入编号、插入时间、删除字符、插入视频时长、插入分辨率、文件名大小写转换、文件名重构、文件真实后缀名还原等操作,可以选择是否按自然数排序,按自然数排序能保证文件顺序是正常顺序即1.MP4,2.MP4,3.MP4而不会出现1.MP4,11.MP4,2.MP4这种情况。如果想要更加强大的功能体验,推荐使用"Bulk Rename Utility"。
- "视频截取":用于截取视频片段。主要应用场景: 1.去除视频片头片尾, 2.提取精彩片段。时间框起始时间和结束时间都支持负值,代表倒数时间。可以批量操作,比如要去除抖音保存视频的最后3秒片尾,只需要在结束时间框输入-3即可,会自动计算对应视频的时间区间而不用挨个设置。支持修改视频帧率。
- **"音频处理"**: 1.从视频中提取音频、从音频中提取音频片段, 2.转换音频格式。时间框也支持负值。支持修改音频采样率、码率。
 - "图片处理": "WEBP转JPG"、"HEIC转JPG"等简单图片类型转换。
- "文本编码转换": 1.批量检测文本编码,导出指定编码文本文件,2.对文本文件的编码格式进行转换,比如批量将'GBK'编码的文本转换为'UTF-8'编码,3.还可以批量转换换行格式,Windows (CRLF)与Unix (LF)相互转换。
- "时间戳操作": 1.转换时间和时间戳, 2.修改文件和文件夹时间戳(支持时间偏移, 支持等差偏移或者随机偏移), 3.检测文件真实数据类型, 4.读取照片EXIF信息、视频、音频元数据信息, 例如: 照片、视频的分辨率以及拍摄时间、经纬度信息、高度信息(若该照片文件有)、拍摄设备, 视频、音频码率、时长等, 5.修改照片、视频的时间戳为照片拍摄时间, 6.导出照片、视频、音频元数据到excel表格。
- "文件分类":识别文件真实数据类型,并分类。支持: 1.按文件类别与数据类型分类, 2.按文件时间(创建时间、修改时间、拍摄时间)分类, 3.按媒体时长分类, 4.按媒体分辨率分类, 5.按拍摄设备信息分类, 6.按EXIF或者视频音频元数据字段分类, 7.按文件大小分类。

程序启动方式一(运行源码):

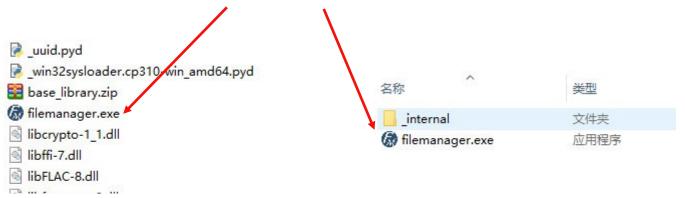
- 1.需要有python环境
- 2.需要安装依赖包

pip install pillow pip install windnd pip install natsort pip install exifread pip install pywin32 pip install pillow_heif pip install filetype pip install chardet pip install pymediainfo



程序启动方式二(直接运行打包好的exe程序):

直接运行解压好的程序目录下的 filemanager. exe 程序





"查找重复文件"功能界面 支持七种检索模式: 1. 同名 2. 相同大小 功能: 查找目录下的重复文件, 并导出至指定目录 3. 相同修改时间 4. 同名且大小相同 5. 同名且修改时间相同 6. 大小相同且修改时间相同 7. 同名且大小相同且修改时间相同 8. 视频时长相同 9. 视频分辨率相同 重复文件的处理方式 10. 图片分辨率相同 **▼** FileManager 操作 设置 查找重复文件 源目录: 导出路径: 查询模式: 同名 操作模式: ● 拷贝 ○ 剪切 ☑ 过滤: ○ 排除 ◎ 选中 (后缀名用逗号','分隔)

勾选<mark>此</mark>选项开启过滤功能,并激活过滤选项框和输入框

支持<mark>文</mark>件过滤,可以选择选中或者排除某些文件类型

还原文件

点此将导出的重复文件

还原回原来位置

X

П

浏览

浏览

执行

点击"执行"按钮开始比对,并会对重

复文件进行操作

注意!

点此查看导出目录

本程序每个功能界面都有按钮或者输入框组件默认锁定,比如本页面的过滤输入框, "还原文件", "撤销还原"按钮,在满足一定条件才会被激活。比如只有勾选了"过滤"的选项框,过滤输入框才会激活可以使用。只有执行了文件查重操作, "还原文件"按钮才会激活可用,只有执行了"还原文件"的操作, "撤销还原"按钮才会激活可用。

点此撤销还原操作

结果显示区

如果过滤多个文件类型, 则每个后缀名用逗号隔开

"查找重复文件(hash模式)"功能界面

功能:查找目录下所有重复文件并导出,比对的是文件的hash值

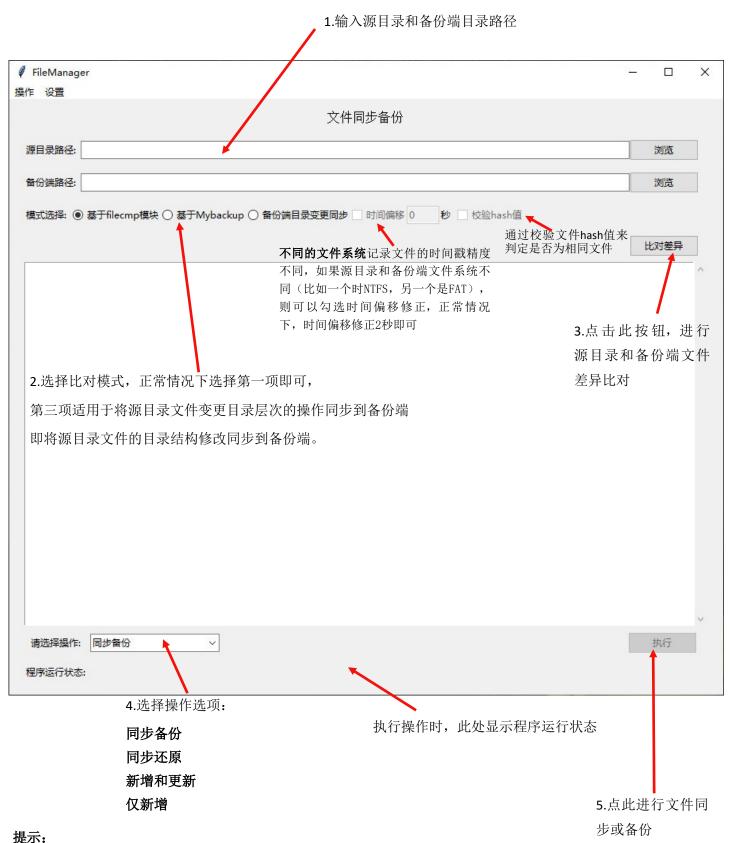
根据文件修改时间顺序对文件进行排序 这个选项框选中,则表示根据 这个选项框选中,则会将所有hash值一样的文件 文件修改时间对文件进行排序 全部导出 **♦** FileManage X 操作 设置 查找重复文件(hash值) 源目录: 浏览 导出路径: 浏览 模式选择: 🗌 根据修改时间排序 🌑 正序 🔾 倒序 文件操作: 📵 拷贝 🔾 剪切 🔲 导出所有hash值相同的文件 执行 执行文件查重操作 重复文件处理方式 进度条 查看重复记录 查看操作失败记录 查看结果 还原文件 撤销还原 如果有操作失败的文件,这个按钮会点亮, 点击即可查看哪些文件操作失败

查看重复记录,查看重 复文件之间的对应关系

"文件同步备份"功能界面

功能: 比对源目录与备份端目录内容,并进行文件同步或备份操作

注意:本功能界面暂不支持文件拖拽自动获取路径功能,源目录路径和备份端路径需要手动输入或者点击浏览按钮选择



涎小:

1.校验hash值耗时很长,通常不需要此项操作

2.文件系统一致时不需要勾选"时间偏移"选项

注意: 只有执行过第3步"比对差异",此处按钮 才会被激活

"还原文件"功能界面

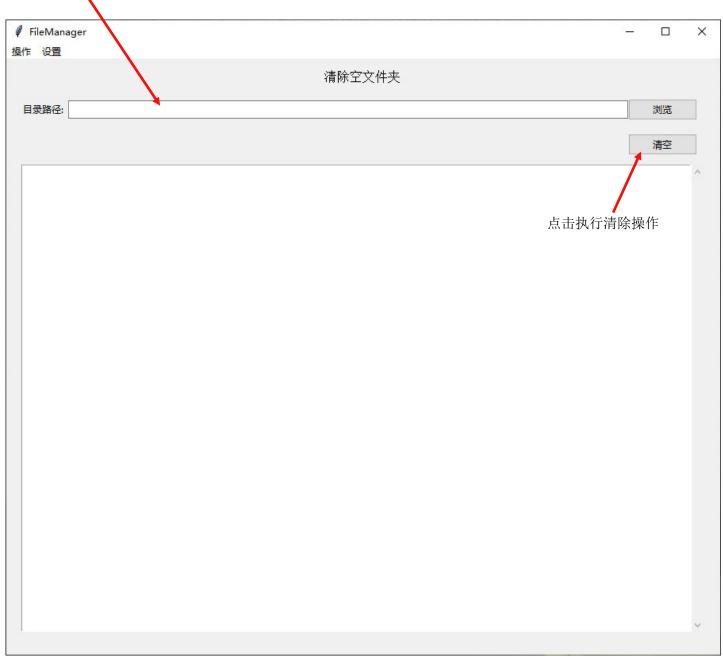
功能:用于根据new_old_record记录还原文件回原来的位置



"清除空文件夹"功能界面

功能:用于清除空文件夹,会递归清除该路径下所有空文件夹,很方便,清除的空文件夹会记录到记录文件中,可以查看

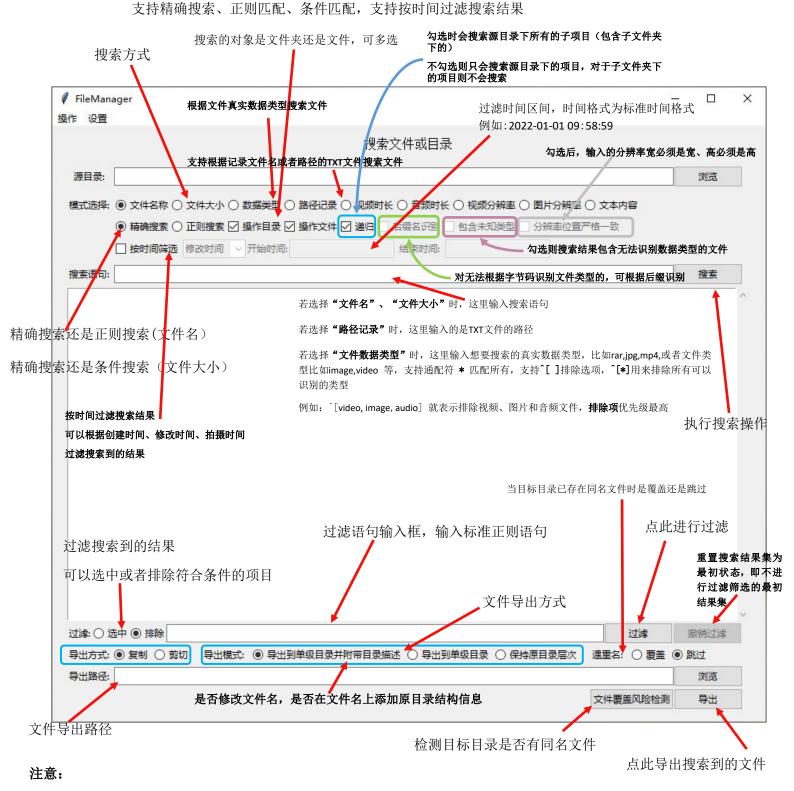
要清除空文件夹的目录路径



"搜索文件"功能界面

功能: 搜索源目录下符合条件的文件、目录, 支持按文件名或者文件大小搜索,

支持按文件真实数据类型搜索,支持根据路径记录搜索,按时长、分辨率、文本内容搜索



1.根据"路径记录"搜索功能模式支持通过记录文件名或者路径的TXT文件内容搜索文件

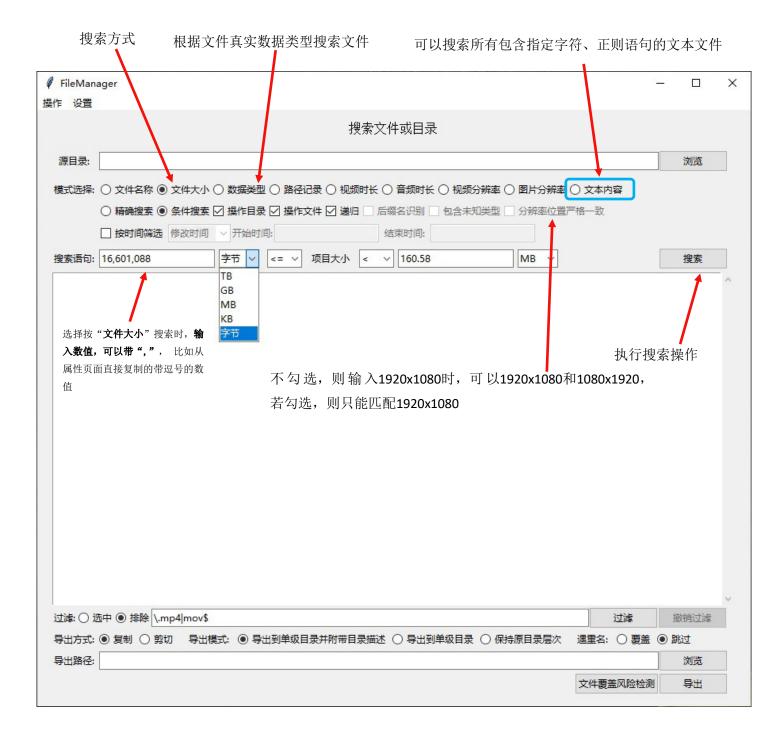
要求:该TXT文件内,每个文件名或路径单独占一行,且该TXT必须为UTF-8编码,且每一行前后无空格支持正则语句,比如 .*jpg .*/bin/.*py

- 2.导出模式支持"导出到单级目录",即会将搜索到的所有文件导出到目标目录下,这种情况是**有文件覆盖风险**的,谨慎选择!
- 比如源目录下有两个文件 a/b.txt和c/b.txt 虽然同名但是父目录不同,但是如果选择"导出到单级目录"就会出现文件重名,会有覆盖风险,导致数据损失!
- 3. "按文件数据类型"搜索时 *^[jpg,video] 表示匹配所有可以识别文件数据类型的文件,然后排除jpg和视频类型的文件

"搜索文件"功能界面

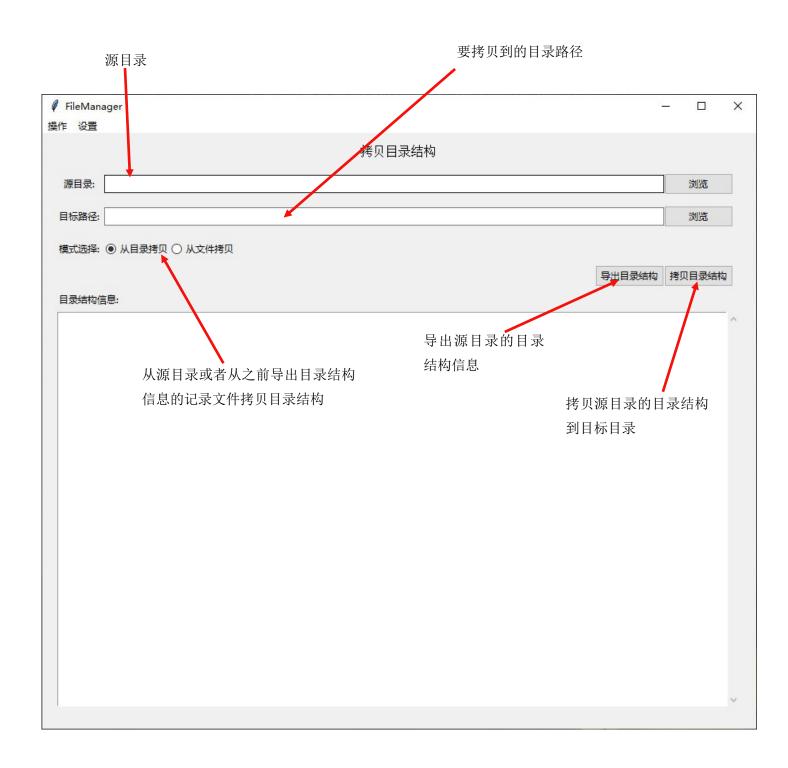
功能:搜索源目录下符合条件的文件、目录,支持按文件名或者文件大小搜索,支持按文件真实数据类型搜索,支持根据路径记录搜索,

支持精确搜索、正则匹配、条件匹配,支持按时间过滤搜索结果



"拷贝目录结构"功能界面

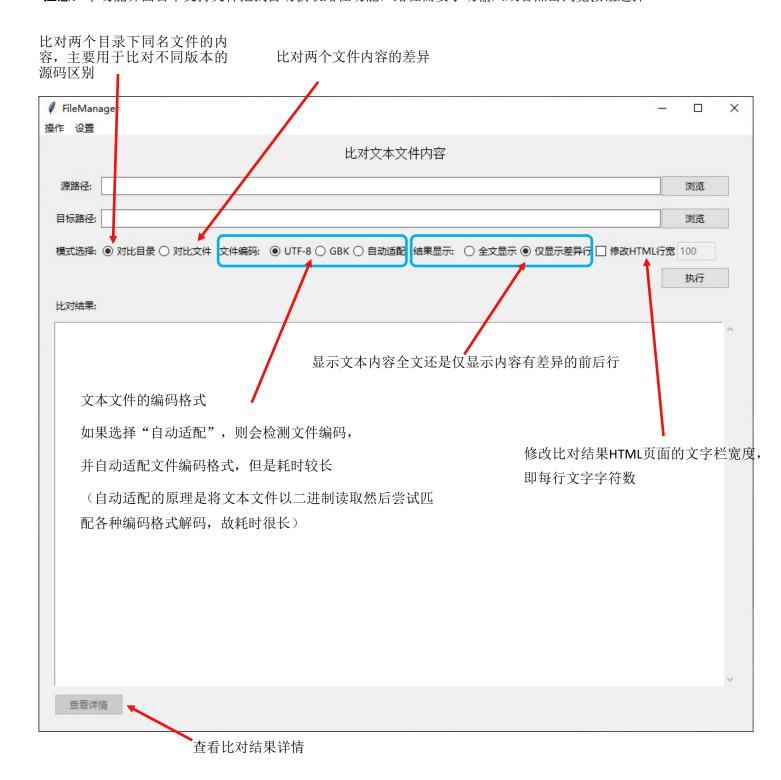
功能: 拷贝或导出源目录的目录结构



"比对文本文件内容"功能界面

功能: 用于比对两个目录下的文本文件内容

注意: 本功能界面暂不支持文件拖拽自动获取路径功能,路径需要手动输入或者点击浏览按钮选择



注意:

- 1.只会比对文本类型文件,包括txt,html,py等文件,不支持比对二进制文件
- 2.选择"比对目录"选项时,两个目录的目录结构要一致,并且文件名相同,这个选项开发出来是用于比对不 同版本的源码区别;

选择"比对文件"选项时,可以选择任意两个文本类型文件进行比对

"计算文件hash值"功能界面

功能:用于计算文件hash值

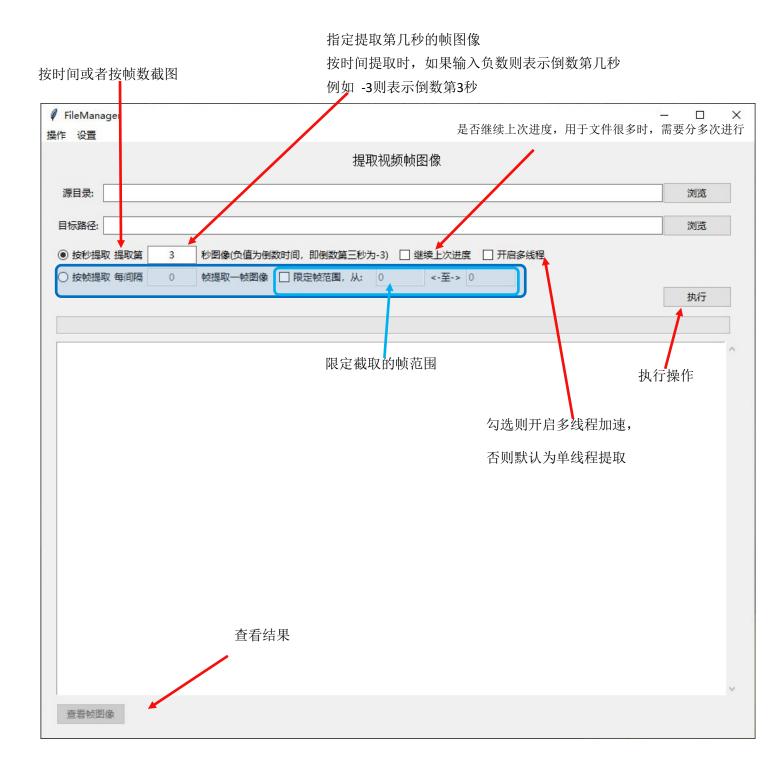


注意:

可以输入目录或者文件路径也可以点击浏览按钮选择,然后点击计算按钮执行计算操作也可以直接拖拽文件或者目录到程序界面,直接计算(推荐)

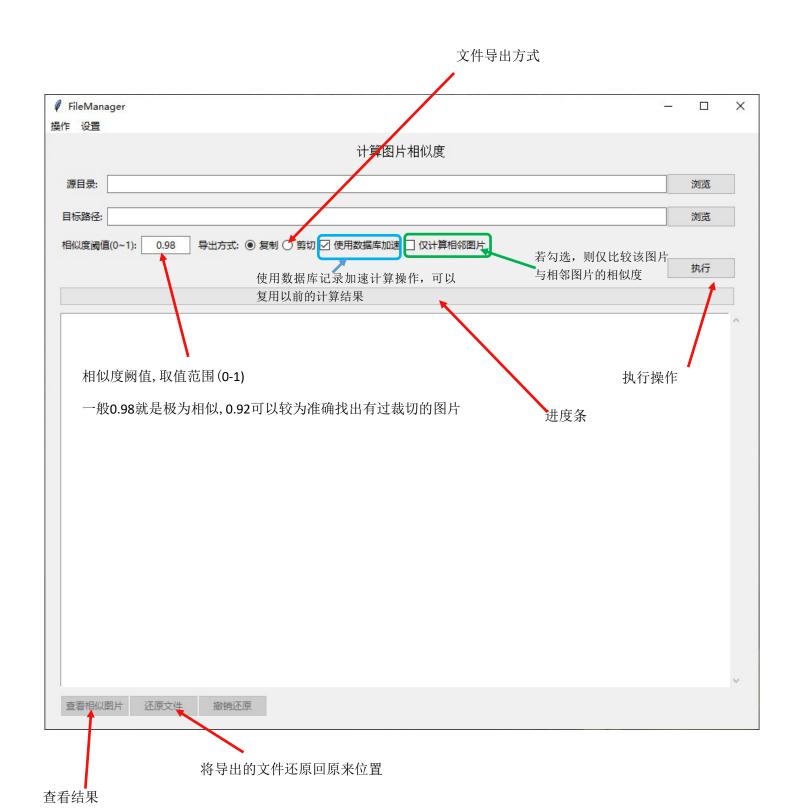
"提取视频帧图像"功能界面

功能: 提取视频的某一帧图像或者某一秒图像



"查找相似图片"功能界面

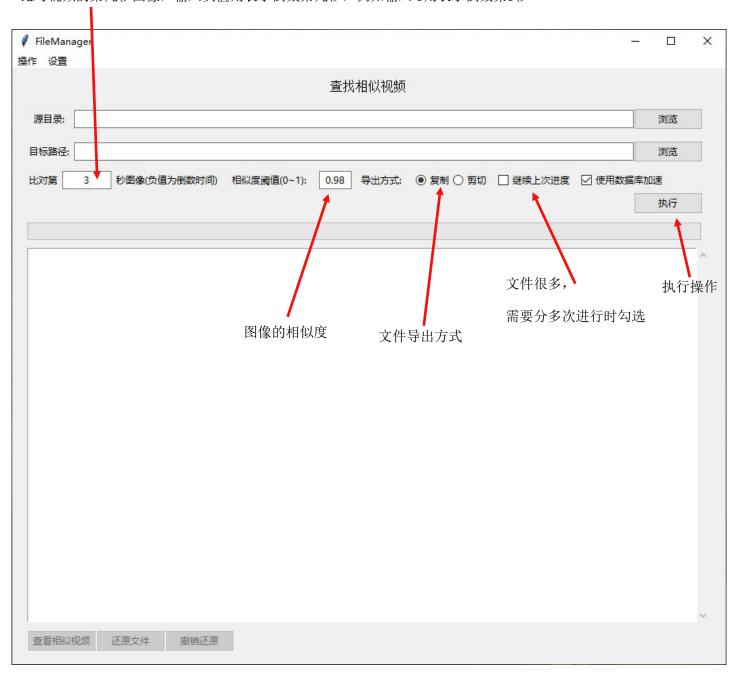
功能: 通过计算图片的相似度, 查找相似图片, 并导出



"查找相似视频"功能界面

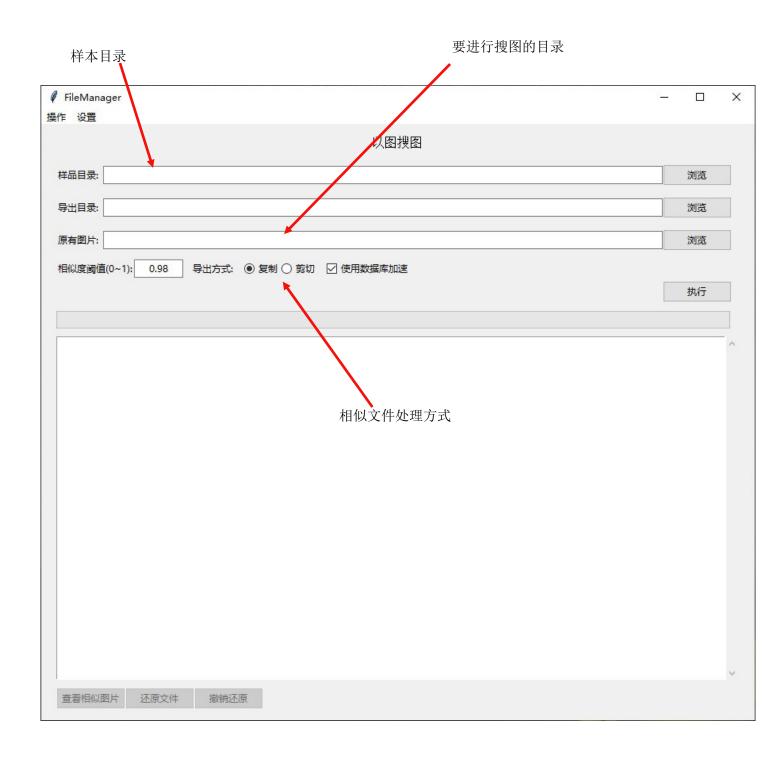
功能:通过比对视频的第几秒图像的相似度,找出相似视频,并导出

比对视频的第几秒图像,输入负值则表示倒数第几秒,例如输入-3则表示倒数第3秒



"以图搜图"功能界面

功能: 通过计算图片的相似度, 实现以图搜图



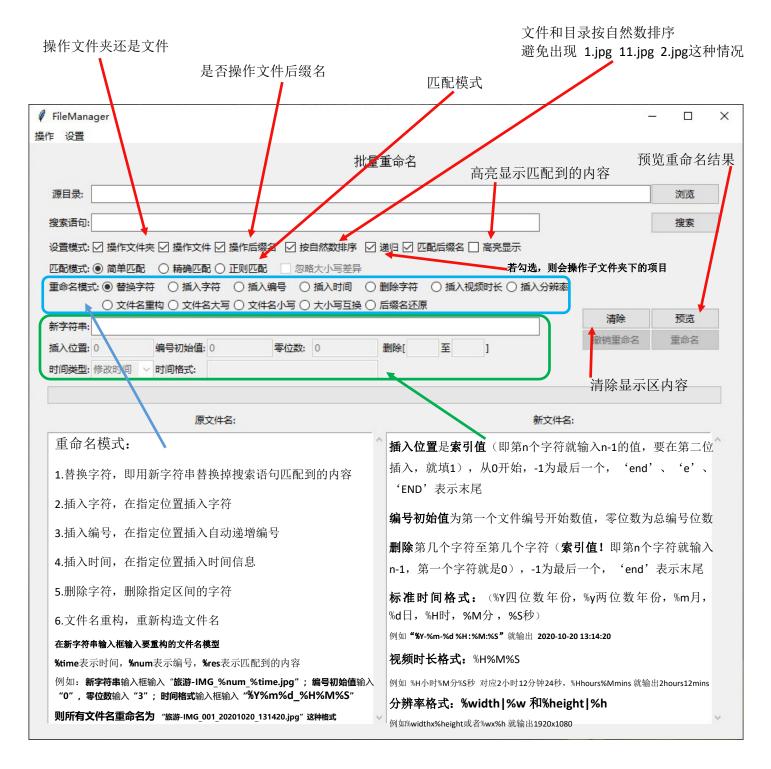
"以视频搜相似视频"功能界面

功能:通过比对视频第几秒的图像相似度找出相似的视频

样本目录	要进行搜索的目录	
		×
	以视频搜索视频	
样本目录:	צועל	
导出目录:	ياريخ ا	ā
原有视频:	يُللان الله الله الله الله الله الله الله	范
比对第 3 秒图像(负值为倒数时间)	相似度阈值(0~1): 0.98 导出方式: ● 复制 ○ 剪切 □ 继续上次进度 ☑ 使用数据库加速	
	执行	Ī
		^
查看相似视频 还原文件 撤销还原		

"批量重命名"功能界面

功能: 批量重命名文件或文件夹



操作顺序:

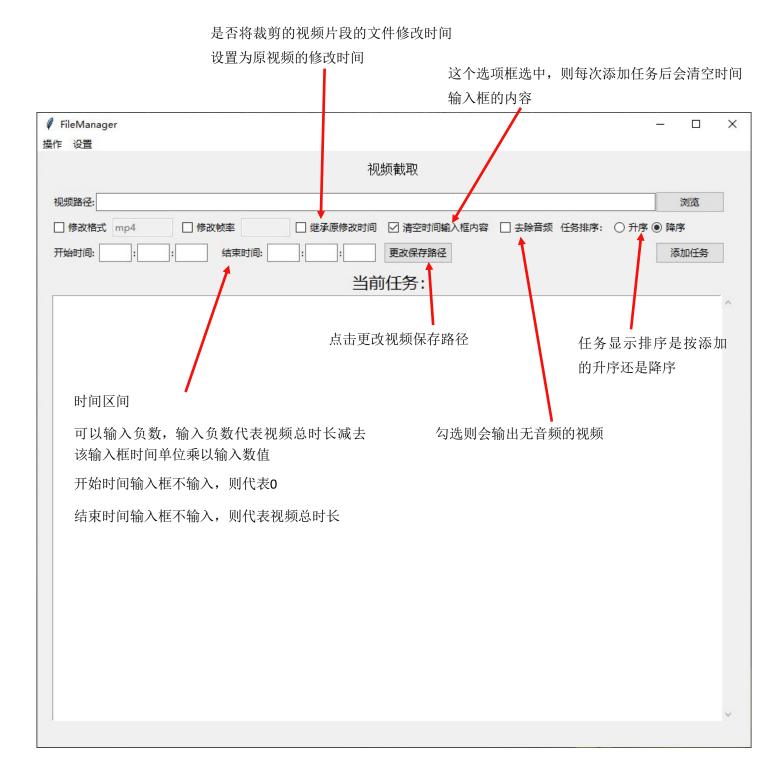
- 1.输入路径、搜索语句,设置操作模式和搜索模式
- 2.点击搜索按钮,搜索符合条件的文件/目录
- 3.选择重命名模式,输入重命名模型或字符串,点击"预览"按钮,预览重 命名效果
- 4.点击重命名按钮,执行重命名操作

注意:

- 1.不勾选"操作后缀名"选项,则只会重命名文件名部分,不会改动原来的后缀名。
- **2.**不勾选"递归"选项,则只会操作源目录下的项目,不会对源目录下子 文件夹下的项目重命名。
- 3.不勾选"匹配后缀名",则只会匹配文件名不含后缀名的那部分,该选项通常和"操作后缀名"同步使用

"视频截取"功能界面

功能: 裁剪视频片段,可以修改视频格式,帧率,去除音频



注意:

若输入的路径是目录或者拖拽文件夹进入程序界面,则会批量操作该目录下的所有视频

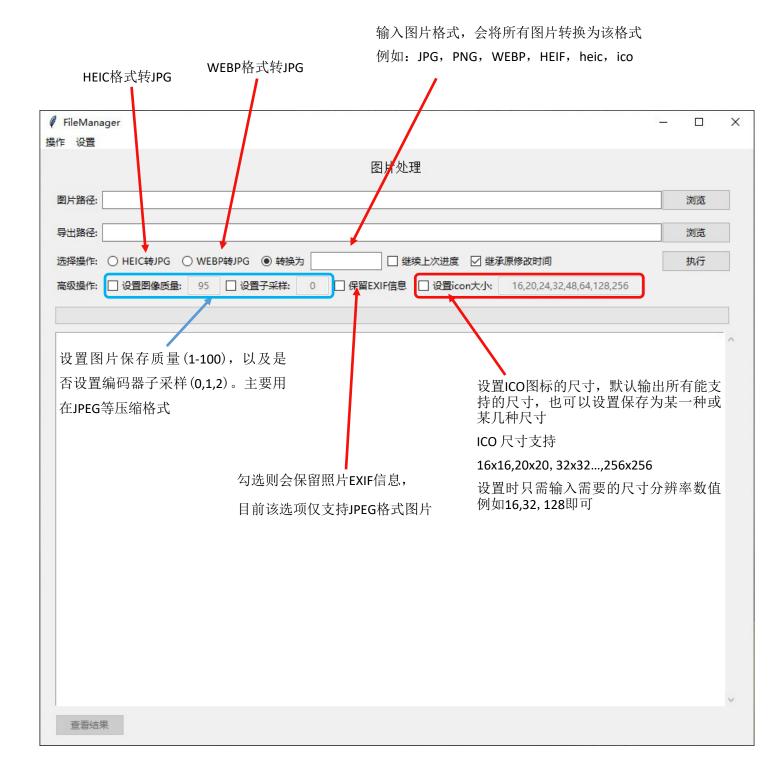
"音频处理"功能界面

功能: 从视频提取音频, 从音频截取音频片段, 转换音频格式

	采用原音频的编码,需注意,只有当 输出的音
勾选则可以修改音频采样率	格式和原音频格式一致的时候才能使用这个选项
默认为44100,即44.1MHZ	勾选则可以修改音频比特率,192k、128k、320k
	×
操作 设置	提取传换音频
源目录:	浏览
导出目录:	浏览
□ 修改格式 □ 修改音频采样率	□ 修改音频比特率 □ 清空时间輸入框内容 □ 采用原编码 ☑ 继承原修改时间
□ 按时间截取,从: : 至:	: 目标路径已存在同名文件: 〇 跳过 〇 覆盖 ⑨ 询问 执行
	
时间区间	当目标目录已存在同名文件时的处理方式
勾选时可以按指定时间区间截取	
可以输入 负数 ,输入 负数代表总时长	咸去 该 输入框时间 单位乘以输入数值
开始时间输入框不输入,则代表0	
结束时间输入框不输入,则代表总时-	K
	~
查看结果	

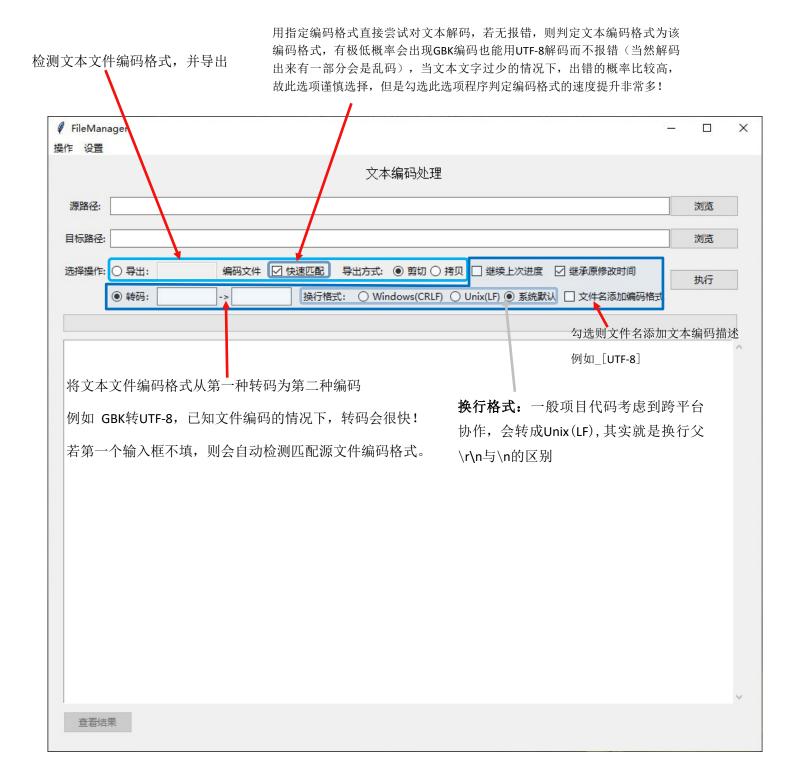
"图片处理"功能界面

功能: 检测图片真实数据类型, WEBP转JPG, HEIC转JPG, 将图片转换为其他格式



"文本编码转换"功能界面

功能: 检测文本文件编码格式,文本文件编码格式转换, CRLF与LF转换

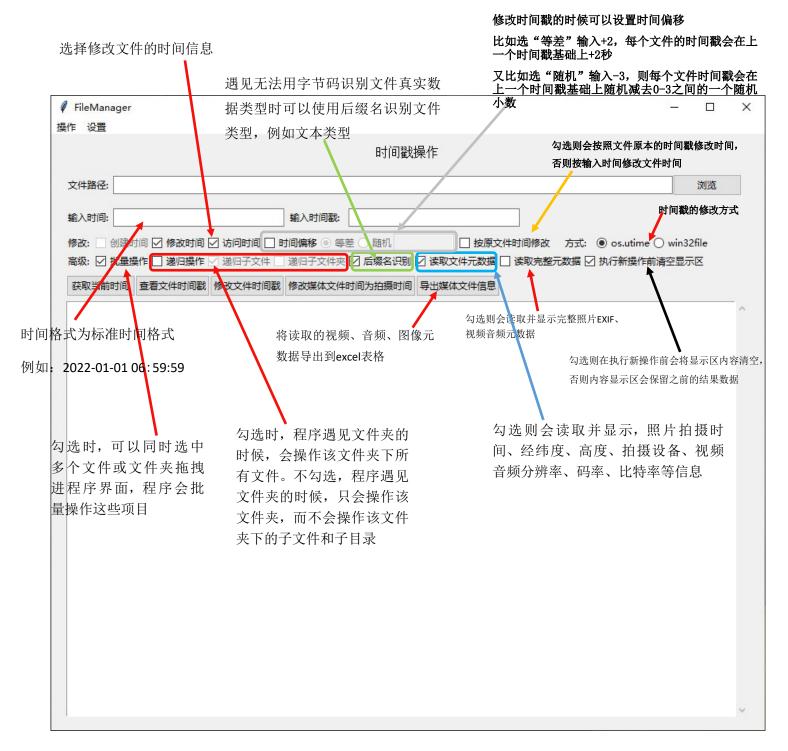


注意:

- 1.检测文本文件编码格式耗时很长,原理是以二进制读取文本文件,然后匹配各种编码格式尝试解码。
- 2.已知源文件编码的情况下直接输入,转码会很快!

"时间戳操作"功能界面

功能:时间和时间戳相互转换,获取文件时间戳,修改文件时间戳,读取照片的拍摄时间,修改照片时间为拍摄时间,读取照片拍摄的经纬度,获取文件真实数据类型



注意:

1.如果文件是照片,则会读取照片的分辨率、拍摄信息(如果该照片文件有EXIF信息),包括拍摄时间,经纬度拍摄设备等信息,

如果是视频、音频则会读取分辨率、帧率、码率、比特率、拍摄设备、经纬度高度等信息

2.修改时间戳的方式:

win32file可以操作文件的创建时间、修改时间、访问时间

os.utime可以操作文件夹的时间但无法修改文件和文件夹的创建时间

- 3.勾选"时间偏移"选项,每个新时间戳都会在基准上加上时间偏移量,偏移量单位为秒数,可接受小数
- 4.勾选"按照原文件时间修改"选项,修改时间戳计算时会以原文件的时间戳问基准计算新时间戳,否则按输入框输入的时间信息为基准计算新时间戳

"文件分类"功能界面

功能:根据文件类别、数据类型、文件大小、拍摄设备、拍摄信息、元数据字段对文件自动分类

件类别: video, image	对于无法用字节码识别数据类	勾选则按MIME类型名创建目录		
据类型: mp4、jpg、text	型的文件用后缀名识别	/		
1				
		/		×
操作 设置				^
JAIP VAL	$\dot{\forall} c b$	EXIF数据段、视频或音频 #分类		
	/ X1*	元数据数据段		
文件目录: C:\Users\Na\Deskt	top\RES2		浏览	
模式选择: ◎ 文件类型 ○ 文件	‡时间 ○ 媒体対长 ○ 媒体分辨率 ○ 设备信息	○ 自定义数据段 ○ 文件大小		
参数选择: 文件类别	✓ ☐ 后缀名识别 ☐ MIME归类 ☐ 分辨	率位置严格一致		
参数设置: 文件 类 别 参数设置: 数据类型			归类	100
双角天王		是否严格要求分辨率位置必须固定	1	
→ 101, pp 1, 300 ← 3kt.				
	7四位数年份,%m月,%d日,%H时,%M分	r),%S秒) 将文件归	1米軟電	
%Y-%m-%d 目录名按2020		村 文 行 炉	1矢登垤	
%Y-%m-%d_%H%M 目录:	S按2020-01-30_0930创建			
视频时长参数: %H %	6M %S 组合			
%H%M%S 目录名为1houri	30min9sec 创建			
%H 则为 1hour 创建				
%H%M 则为 1hour30min 包	 			
7011701VI X373 IIIOUI 3011IIII 2)Æ			
	%GB %KB %B 或者%M %G %K 组合			
%MB 则目录名按 120.2MB 6 %G 则目录名按1.3GB 创建	则 建			
加州日水石1以1.500 凹座				
分辨率参数	、 %height 或%w、 %h分别表示宽、高	Ŧ Ŧ		
77//1/2/201		•		J
文件操作模式: ● 导出到单级目	目录并附带目录描述 () 导出到单级目录 () 保持	原目录层次 選重名: ○ 覆盖 ® 跳过		
查看结果 还原文件	撤销还原			

注意:

MIME类型数据类型名和我们常用的数据类型名可能会有些不一样

程序设置界面

功能:设置程序设置信息

FileManager : 设置			8_)
	设置	置界面			
保存记录的目录:	C:\Users\Na\Desktop\Projects\FileManager\Record	Н		查看	
保存数据库的目录:	: C:\Users\Na\Desktop\Projects\FileManager\db			查看	
保存日志的目录:	C:\Users\Na\Desktop\Projects\FileManager\log			查看	
FFMPEG路径:	C:\Users\Na\Desktop\Projects\FileManager\imageio_ffmpeg\binaries\ffmpeg-win64-v4.2.2.exe			查看	
SAFE_DIR目录名:	SAFE_DEL				
系统的编码格式:	GBK	受磁盘IO限制,IO线程数不建	议超过	5,	
IO线程数:	4	计算线程数也不建议开启过多			
计算线程数:	4				
勾选时,册	® 使用安全删除 遇到路径 使用安全删除 遇到路径 除文件不会直接从硬盘删除,而是会算恢复(类似系统的回收站一样),防止			夏默认设置	ghri
		1.激活修改权限			

注意:

设置界面,默认所有选项是不可操作,要想修改程序设置,需先进行"权限验证",修改后要点击"保存到设置"按钮,方才修改生效