

利用 Excel VBA 实现数据自动抽取汇总

作者：徐宝平 日期：2011 年 12 月 26 日 公司：MTIW

前言：

对于电子工程师和测试工程师来说，最头疼的莫过于整理测试了半天得到的数据，这些数据大部分是由软件自动测试生成并保存为.txt .log 等类似的文本格式。这些文本中的数据的分布的格式及文本的整体框架是一致的，这就为我们自动整理这些文本文件提供了条件。

下面我要介绍一款软件，它是基于 Excel 和其中的 VBA 语言进行对文本格式的数据自动抽取和整理，并保存成我们自己定义的格式，前期介绍的是一维的格式数据整理，能将 200 多个文件的数据整理到一张 sheet 中。以便工程师分析和对比在不同条件下的前后数据的差异，从而找到问题的关键。

第一章 软件界面的介绍

这款软件的名称叫 **标准相同格式文本整理-共享版.xls** 之所以叫做共享版，是因为这是原软件的删改版，原软件是为我们公司特定的数据格式而编写的，能粗略的自动删改并对齐数据的格式，而这款软件只能对完全相同的数据格式的文本文件进行处理。如果大家有基本上格式相同的文本文件需要处理，可以发我，尽力修改成为您特制的数据整理软件。我的 QQ：751623467，加好友的时候注明“数据整理”。不过我一般不怎么上线，那就发我邮箱吧！

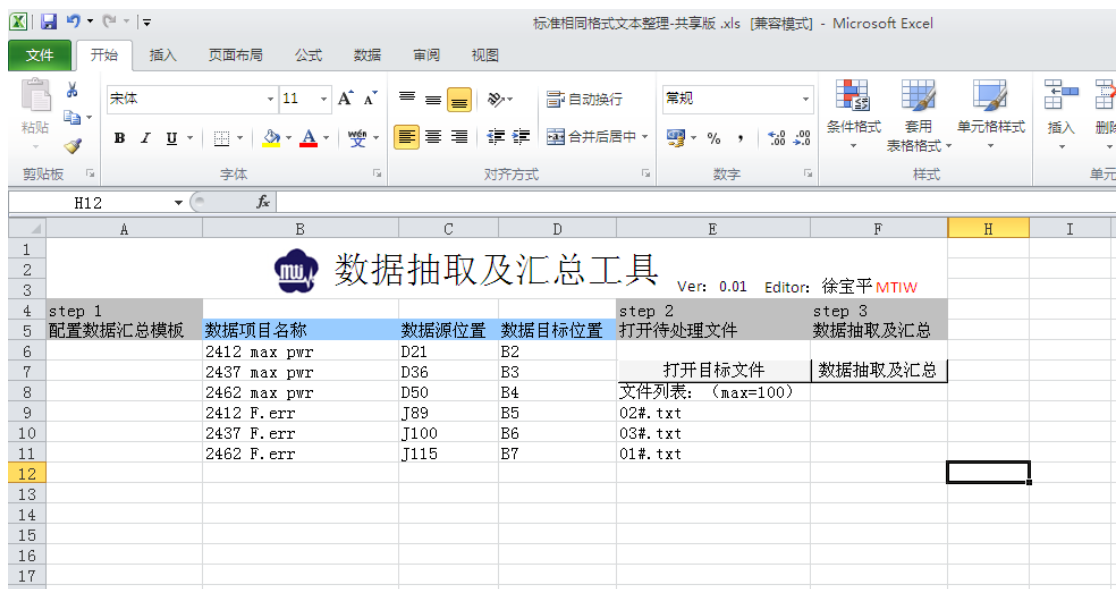


图 1



上图是软件的主操作界面，其中 **江扬科技(无锡)有限公司** 是我所在公司的 LOGO，因为这款软件的诞生的前提是我在这家公司工作，舒适的工作环境及友善的上司使得这款软

件能面世。做出这个软件的初衷是因为我的工作需要，公司是做无线射频，部门是做 WIFI 产品，用的是 IQ 设备和配套软件，用这个软件测试 AP 的时候会得到大量的.log 文件。平时对产品的分析及改进也主要集中在对这些 log 文件的分析，改变电路时的不同 log 文件的对比分析。以前整理这些数据的时候都是手动复制粘贴到 excel 中，然后才能分析，这会花费大量的时间和精力。整理 10 个文件中的近 100 项数据就得复制粘贴 1000 多次，难免会发生错误或是产生厌烦的心理。这样的环境催生了这款自动数据抽取整理软件的诞生。

这款软件的界面主要分成 3 个部分。第一部分是**配置数据汇总模板**及其后面的 3 列。B 列是数据项目名称（最多支持 1000 项数据）。C 列是数据尤其关键，它是这个软件从原始文本中读取数据的关键依据，怎么得到这一列的数据，下面我会详细的介绍。D 列是数据目标位置，它是标志将从原始文本中得到的数据放入目标 sheet 中的位置。

数据项目名称	数据源位置	数据目标位置
2412 max pwr	D21	B2
2437 max pwr	D36	B3
2462 max pwr	D50	B4
2412 F. err	J89	B5
2437 F. err	J100	B6
2462 F. err	J115	B7

表 1

下面示例，表 1 是对 6 项数据进行整理的步骤一，这时的软件只整理 6 项数据（最多能整理 1000 项数据，理论上能整理 60000 多项，但考虑到大家的数据项一般没有这么多，为了节约系统的开销，暂定为 1000 项），D21 是数据在 log 中的位置，B2 是将要放到整理好的 sheet 中的位置。

下面重点讲一下怎么得到数据源位置（C 列）。

Setp1: 任意打开或新建一个 excel 文件，打开文件

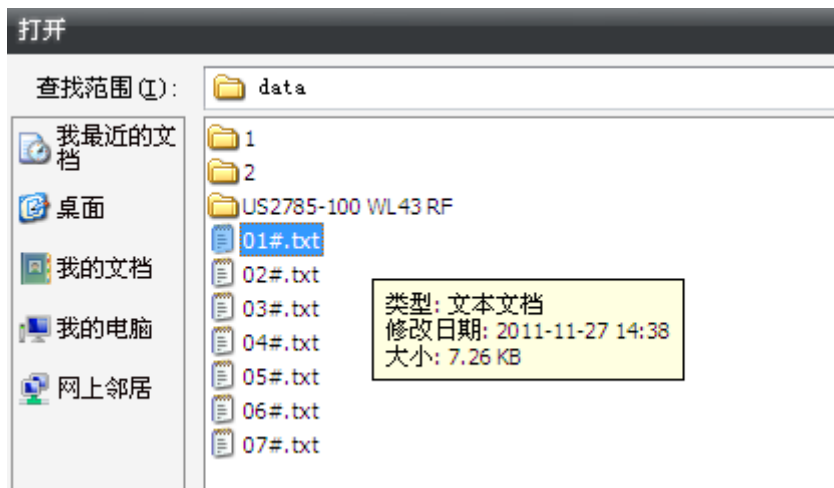
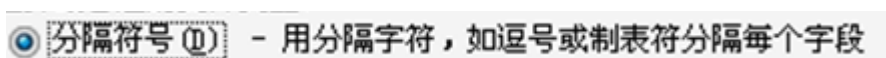


图 2

选择需要整理的数据格式，我选择打开上图中的 01#.txt 文件，目前支持的格式包括 .log .txt .dat 等文本文件。

Setp2: 出现下图所示的对话框，提示文本导入向导，默认选择分隔符号



点击下一步。

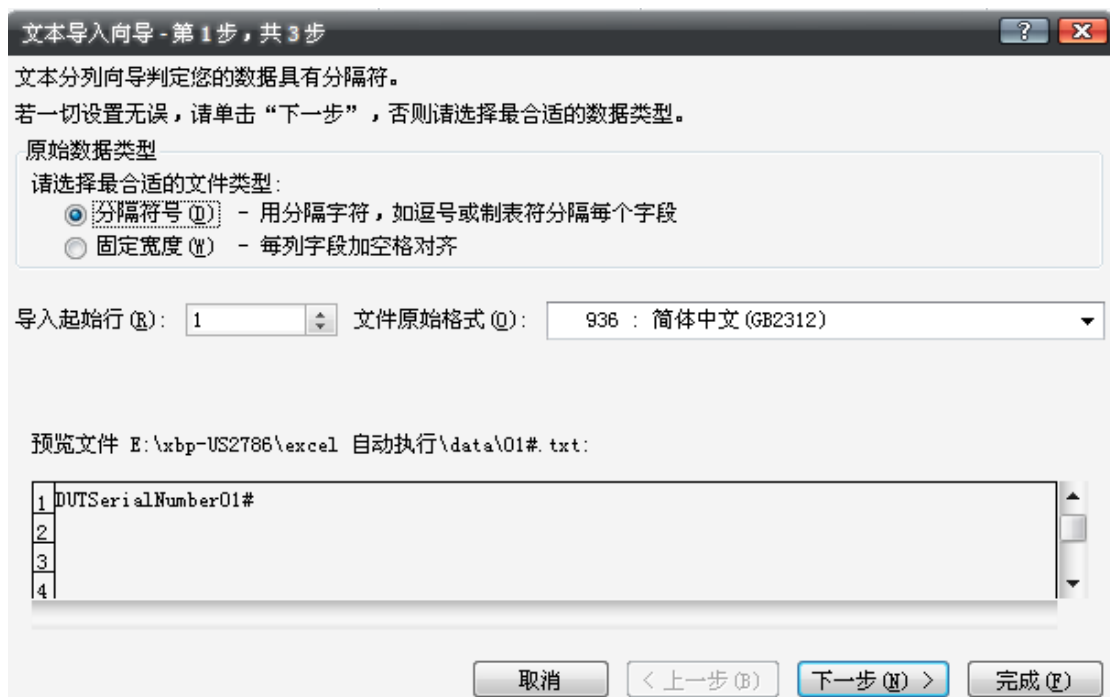
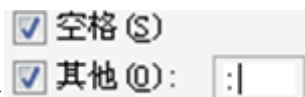


图 3

Step 3:

文本导入第二步, 按图 4 所示的标志选择勾选空格和其他中填入 英制的冒号 : 点击下一步。



和其他中

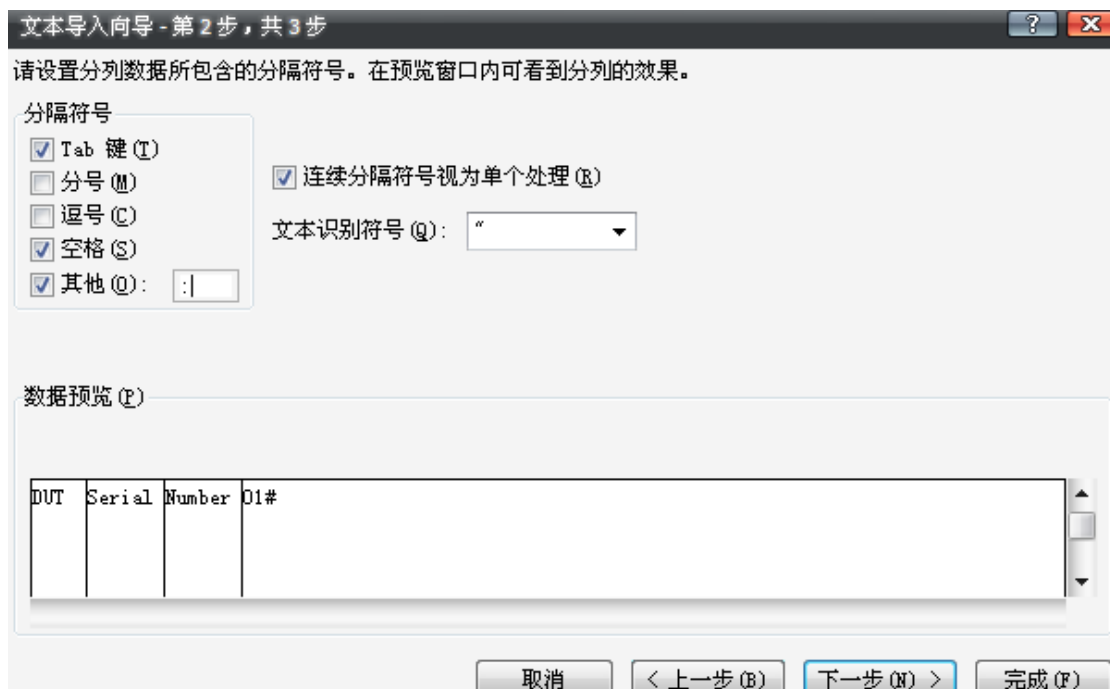


图 4

Step 4:

得到下图所示的文本导入向导 第 3 步, 不做任何动作, 点击完成按钮。

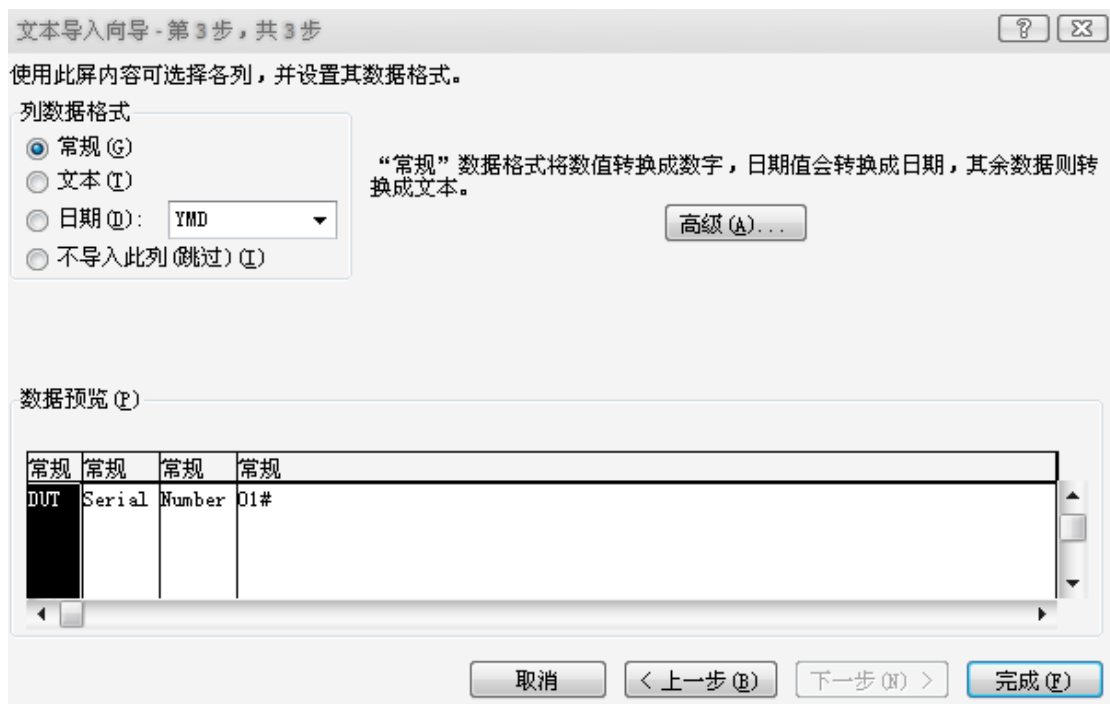


图 5

Step 5

得到下图所示的文件

	A	B	C	D	E	F	G
1	DUT	Serial	Number	01#			
2							
3							
4							
5		Calibratiis		being	performedby		
6		-	Revision	0.9	BUILD	#34	ART_11n
7		-	Customer	Version	(ANWI	BUILD)-	
8	Devlib	Revision	0.9	BUILD	#33...	wxcART_11n	
9		Using	config	file	//depot/bringup/ar5k/config		
10							
11	RX	main	ANTENNA	=	A		
12							
13	using	cal	fixed	gain	for	mode	11g
14							
15	*****						
16	*	Manufactu	Calibrati	Start	*		
17	*****						
18							
19	Collectirraw		data	for	the	adapter	for
20	ch	2412	-->	pcdac	72		
21	max	pwr	is	22	dBm		
22	pcdac	22	pwr	0.49	max	22	
23	pcdac	42	pwr	8.47			
24	pcdac	15	pwr	1.51			
25	pcdac	25	pwr	2.04			
26	pcdac	36	pwr	5.05			
27	pcdac	46	pwr	9.94			

图 6

出现了数据整理项目的第一个数据位置 D21, 这份 sheet 中的你所需要的数据的位置就是你
需要填在 C 列的数据位置。逐个将需要整理的数据位置填入到 C 列中。

下面介绍一下 D 列的数据，这个就是数据整理好后的位置。整理好得数据将放到 sheet2 中。

注意：整个软件一共有 3 个 sheet，sheet1 是软件的主操作界面，sheet2 是数据整理好后的保存的位置，sheet3 是对原始数据的暂放位置，这 3 个 sheet 不能更改名称也不能删除任何一个，不然程序会报错。

第 2 章 软件的使用

在进行数据操作之前，需要将待整理的项目、源地址及目标地址的 3 列填好

软件一共有两个按钮，一个是 **打开目标文件**，一个是 **数据抽取及汇总**，

第一个按钮是用来打开待整理的文件的，支持多选，但待选的多个文件只能在同一个文件夹内。

单击打开目标文件按钮，出现下图，选中需要整理的文件，单击打开按钮。

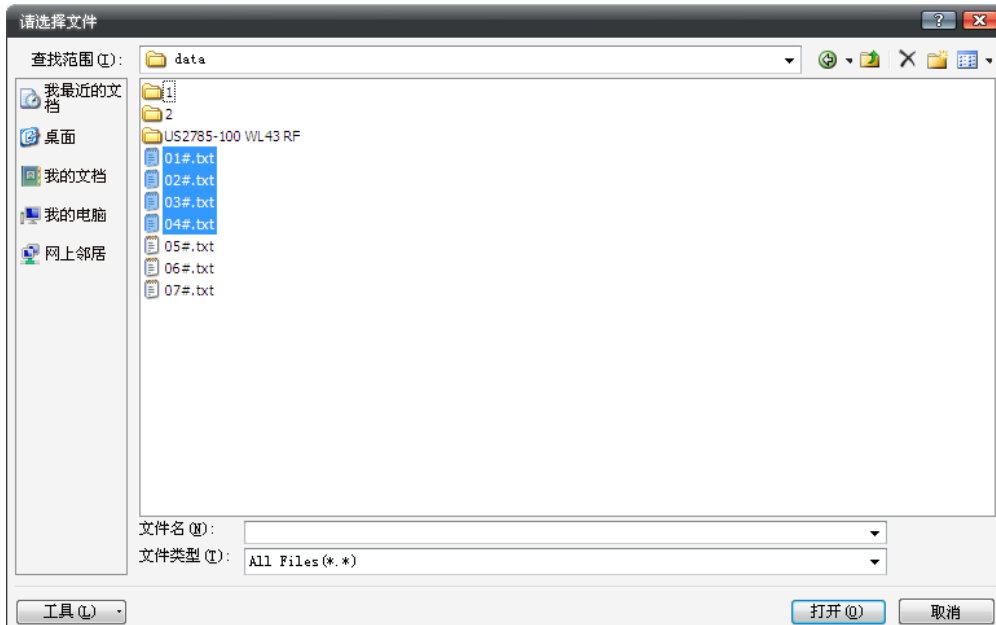


图 7

刚刚选中的 4 个文件会出现在如下图所示的位置。重新确认无误后，单击数据抽取及汇总按钮，若选中的文件有误则需要重新选择文件，该软件不支持增加文件，不支持删除部分文件。后续需要改进的地方，呵呵！

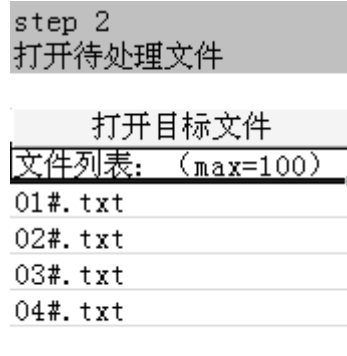


图 8

点击数据抽取及汇总按钮，则出现对话框如右图，告知您一共待处理的文件个数，待确认后单击**确认**按钮，则软件就会快速执行抽取汇总任务，汇总结束后回跳出另存为对话框，默认的文件名是软件 excel 的名字，如需要保存则在改名字后按**保存**按钮，如不需要保存则单击**取消**按钮。

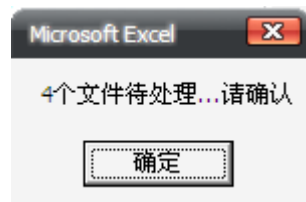


图 9

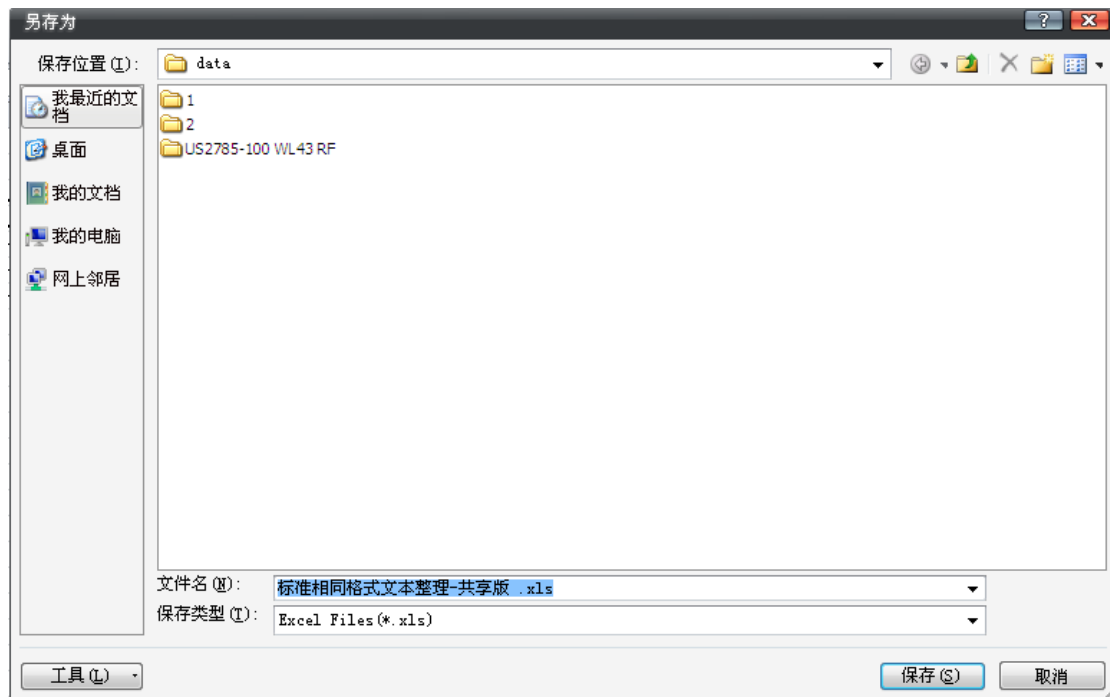


图 10

若选择取消，则会自动跳到 sheet2 中显示已经整理好得表格，看一下效果吧，还不错吧，呵呵！

	A	B	C	D	E	F
1	S/N	01#	02#	03#	04#	
2	2412 max pwr	22	22	26	23	
3	2437 max pwr	22.05	22.05	25.05	24	
4	2462 max pwr	20.96	20.96	20.96	25	
5	2412 F. err	0.5ppm	1.5ppm	0.5ppm	0.5ppm	
6	2437 F. err	0.4ppm	2.4ppm	0.4ppm	0.4ppm	
7	2462 F. err	0.4ppm	0.4ppm	0.4ppm	0.4ppm	
8						
9						
10						
11						

图 11

这就是将 4 个文件中的 6 项数据从原始文档中提取处理并放到一个 sheet 中进行对比分析的自动数据抽取汇总软件的基本功能。

想得到这款软件就到百度文库中搜索 **标准相同格式文本整理-共享版.xls** 就能得到啦，免费软件，欢迎使用和改进。

第 3 章 源程序的讲解

为了让大家能充分使用和理解这款软件现将整个程序的代码附上,包含了相当充分的注释,相信只要是有心人,一定能打造属于自己的数据自动整理软件。

```
Private Declare Function apiOpenClipboard Lib "user32" Alias "OpenClipboard"  
(ByVal hwnd As Long) As Long  
Private Declare Function apiEmptyClipboard Lib "user32" Alias "EmptyClipboard" ()  
As Long  
Private Declare Function apiCloseClipboard Lib "user32" Alias "CloseClipboard" () As  
Long  
Sub action_Click()  
    Dim filetoopen(100) As String '定义数组用于存放文件的路径 大小为 100  
    Dim filename(100) As String '用于存放文件名不包含后缀名  
    Dim filenamesuffix(100) As String '用于存放文件名及后缀  
    Dim intl As Integer '用于数组的循环  
    Dim totall As Integer '用于数组的大小  
    Dim totallist As Integer '用于数组的大小  
    Dim nextRow As String 'next row index  
    Dim MyArray() As String '定义动态数组  
    Dim arrayl As Integer '用于数组的大小 行  
    Dim arrayJ As Integer '用于数组的大小 列  
    'Dim fileSaveName1 As String '用于存放文件名  
    Dim xbp As String '临时用  
    Dim softname As String '存放运行程序的文件名  
    nextRow = 0  
  
    softname = ActiveWorkbook.Name  
    Windows(softname).Activate '将窗口提到 z-次序的最前面  
    Worksheets("Sheet2").Activate '激活 sheet2  
    ActiveSheet.UsedRange.Delete '删除 sheet2 中已使用单元格中的数值  
    Worksheets("Sheet1").Activate '激活 sheet1  
  
    totallist = Application.WorksheetFunction.CountA(Range("B6:B1006")) '得到需要  
整理项的个数,最大为 200 项  
    ReDim MyArray(totallist, 3) '重新定义动态数组的大小及维数  
    For arrayJ = 0 To 1  
        For arrayl = 0 To totallist - 1  
            MyArray(arrayl, arrayJ) =  
Worksheets("Sheet1").Range("C6").Offset(arrayl, arrayJ).Value  
        Next arrayl  
    Next arrayJ  
  
    Range("B6", Range("B6").End(xlDown)).Select '选择 B6 及以下全部有值的
```


单元格

```
Selection.Copy
Sheets("Sheet2").Select
Range("A2").Select
ActiveSheet.Paste
Range("A1").Select
ActiveCell.FormulaR1C1 = "S/N"
Selection.HorizontalAlignment = xlCenter    '居中

Sheets("Sheet1").Select
Range("G9").Select
total = Application.WorksheetFunction.CountA(Range("G9:G109")) '统计第
7 行第 9 列开始到结束的有值单元格格式
MsgBox total & "个文件待处理...请确认"

For int1 = 0 To total - 1
    Windows(sofname).Activate '将窗口提到 z-次序的最前面
    Sheets("Sheet1").Select
    filetoopen(int1) = Worksheets("Sheet1").Range("G9").Offset(nextRow,
0).Value '将文件及路径放到数组 filetoopen 中
    filename(int1) = Mid(filetoopen(int1), InStrRev(filetoopen(int1), "\", ,
vbTextCompare) + 1, _
    InStrRev(filetoopen(int1), ".", , vbTextCompare) -
InStrRev(filetoopen(int1), "\", , vbTextCompare) - 1)
    filenamesuffix(int1) = Mid(filetoopen(int1), InStrRev(filetoopen(int1), "\", ,
vbTextCompare) + 1)
    nextRow = nextRow + 1
Next int1

nextRow = 0
For int1 = 0 To total - 1

    Workbooks.OpenText    filename:=filetoopen(int1),    Origin:=936,
StartRow:=1 _
    , DataType:=xlDelimited, TextQualifier:=xlDoubleQuote, _
    ConsecutiveDelimiter:=True,    Tab:=True,    Semicolon:=False,
Comma:=False, _
    Space:=True, Other:=True, OtherChar:=":", FieldInfo:=Array(Array(1, 1),
_
    Array(2, 1), Array(3, 1), Array(4, 1)), TrailingMinusNumbers:=True '打
开待处理的文件

    Cells.Select    '选中打开的待处理文件中的所有单元格
    Selection.Copy    '复制选中内容
```

```

Windows(softname).Activate '将窗口提到 z-次序的最前面
Sheets("Sheet3").Select
Range("A1").Select
ActiveSheet.Paste
Application.CutCopyMode = False
Range("A1").Select
Windows(filename&suffix(intl)).Activate
ActiveWindow.Close savechanges = fail '关闭活动的工作簿 不保存
Call myClr '清空剪切板 以释放内存

```

```

Windows(softname).Activate '将窗口提到 z-次序的最前面

```

```

'执行关键的汇总功能

```

```

Sheets("Sheet3").Select
For arrayl = 0 To totallist - 1 '从 sheet3 中指定位置（由数组的第 1 列决定）
    读出数据并放到数组的第 3 列
        MyArray(arrayl, 2) = Worksheets("Sheet3").Range(MyArray(arrayl,
0)).Value

```

```

Next arrayl

```

```

Sheets("Sheet2").Select
For arrayl = 0 To totallist - 1 '将数组的第 3 列的数值放到 sheet2 的指定
    （由数组的第 2 列决定）位置
        Worksheets("Sheet2").Range(MyArray(arrayl, 1)).Offset(0,
nextRow).Value = MyArray(arrayl, 2)
        Next arrayl
        Range("B1").Offset(0, nextRow).Value = filename(intl)

```

```

nextRow = nextRow + 1

```

```

Next intl

```

```

Worksheets("Sheet3").Activate '激活 sheet3
ActiveSheet.UsedRange.Delete '删除 sheet3 中已使用单元格中的数值
Worksheets("Sheet2").Activate '激活 sheet2

```

```

'保存整理好的工作表

```

```

filesavename = Application.GetSaveAsFilename(fileFilter:="Excel Files(*.xls),
*.xl*", FilterIndex:=1, Title:="另存为")

```

```

If filesavename <> False Then

```

```

    Sheet2.Activate

```

```

    ActiveSheet.Copy

```

```

    'MsgBox filesavename '显示待存入的文件名及路径 调试用

```

```

    ActiveWorkbook.Close SaveChanges:=True, filename:=filesavename

```

'新表关闭

```
        Sheet2.Select
    End If
End Sub

Sub open_sourcefile_Click()

    Dim nfile As String
    Dim nfiletemp As String
    Dim nextRow As String 'next row index
    Dim Filter As String '用于定义打开文件的类型
    Dim filetoopen '定义为数组，存放需要打开的文件名
    Dim intl As Integer '用于数组的循环
    Dim total As Integer '用于存储数组的大小
    Dim lastl As Integer '用于文件名开始的位置在字符串中起始位置
    nextRow = 0 '定义下次的偏移大小
    total = 0

    Filter = "All Files(*.*),*.*,Word Documents(*.do*),*.do*," & _
        "Text Files(*.txt),*.txt,Excel Files(*.xl*),*.xl*"

    filetoopen = Application.GetOpenFilename(fileFilter:=Filter, FilterIndex:=1, Title:="请选择文件" _
        , MultiSelect:=True)

    If Not IsArray(filetoopen) Then

        MsgBox "你没有选择文件", vbOKOnly, "提示"
        For intl = 1 To 200 '将 E9 向下的 200 个单元格内容清空
            Worksheets("Sheet1").Range("E9").Offset(nextRow, 0).Value = Null
            Worksheets("Sheet1").Range("G9").Offset(nextRow, 0).Value = Null
            nextRow = nextRow + 1
        Next intl
        nextRow = 0 '初始化偏移量
    Else
        With Worksheets("Sheet1").Range("E9")
            For intl = 1 To 200 '将 E9 及 G9 向下的 200 个单元格内容清空
                .Offset(nextRow, 0).Value = Null
                .Offset(nextRow, 2).Value = Null
                nextRow = nextRow + 1
            Next intl
            nextRow = 0 '初始化偏移量
            total = UBound(filetoopen, 1) '得到 FileToOpen 中的一维数组的大小
            For intl = 1 To total
                nfile = filetoopen(intl)
```

```
        nfiletemp = Mid(nfile, InStrRev(nfile, "\", , vbTextCompare) + 1) '在包含  
文件路径的字符串中提取文件名
```

```
        .Offset(nextRow, 0).Value = nfiletemp
```

```
        .Offset(nextRow, 2).Value = nfile
```

```
        nextRow = nextRow + 1
```

```
    Next int1
```

```
End With
```

```
End If
```

```
End Sub
```

```
Sub myClr()
```

```
    apiOpenClipboard (0) '打开剪切板
```

```
    apiEmptyClipboard '清空剪切板
```

```
    apiCloseClipboard '关闭剪切板
```

```
End Sub
```