(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2016 年 12 月 22 日 (22.12.2016)

WIPO | PCT

(10) 国际公布号
WO 2016/202308 A3

(51) 国际专利分类号:
G06F 17/30 (2006.01)

(21) 国际申请号: PCT/CN2016/094954

(22) 国际申请日: 2016 年 8 月 12 日 (12.08.2016)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201510329296.6 2015 年 6 月 15 日 (15.06.2015) CN

(71) 申请人: **北京金山安全软件有限公司 (BEIJING KINGSOFT INTERNET SECURITY SOFTWARE CO., LTD.)** [CN/CN]; 中国北京市海淀区小营西路 33 号二层东区, Beijing 100085 (CN)。

(72) 发明人: **焦国强 (JIAO, Guoqiang)**; 中国北京市朝阳区朝阳北路 237 号复星国际中心 12 层, Beijing 100020 (CN)。

(74) 代理人: **北京柏杉松知识产权代理事务所（普通合伙） (PATENTSINO IP FIRM)**; 中国北京市西城区北三环中路 27 号商房大厦 413 室, Beijing 100029 (CN)。

(81) 指定国 (除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明，要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

— 包括关于请求恢复一项或多项优先权要求的信息(细则 26 之二.3 和 48.2(b)(vii))。

(88) 国际检索报告公布日期: 2017 年 2 月 9 日

(54) **Title**: METHODS AND DEVICES FOR FILE NAME IDENTIFICATION AND FILE CLEANING

(54) **发明名称** : 一种文件名称识别及文件清理方法及装置

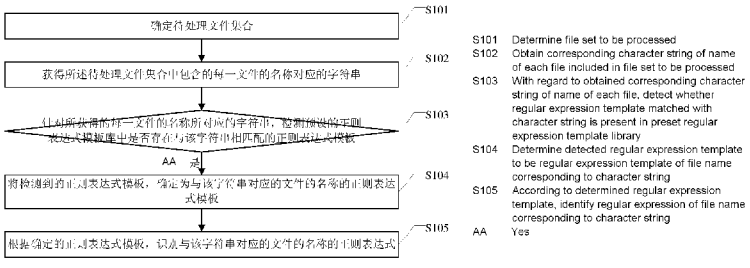


图 1

(57) **Abstract**: Disclosed are methods and devices for file name identification and file cleaning, the method for file name identification comprising: determining a file set to be processed; obtaining a corresponding character string of a name of each file included in the file set to be processed; with regard to the obtained corresponding character string of the name of each file, detecting whether a regular expression template matched with the character string is present in a preset regular expression template library; determining a detected regular expression template to be a regular expression template of the file name corresponding to the character string; according to the determined regular expression template, identifying a regular expression of the file name corresponding to the character string. The present invention implements the automatic identification of file names and the cleaning of files in electronic devices, increasing the efficiency of identifying file names in electronic devices and the efficiency of file cleaning.

(57) **摘要**: 本申请实施例公开了一种文件名称识别及文件清理方法及装置，文件名称识别方法包括：确定待处理文件集合；获得待处理文件集合中包含的每一文件的名称对应的字符串；针对所获得的每一文件的名称所对应的字符串，检测预设的正则表达式模板库中是否存在与该字符串相匹配的正则表达式模板；将检测到的正则表达式模板，确定为与该字符串对应的文件的名称的正则表达式模板；根据确定的正则表达式模板，识别与该字符串对应的文件的名称的正则表达式。应用本申请实施例，实现了电子设备中的文件名称的自动识别及文件的清理，提高了识别电子设备中的文件名称的效率及文件清理的效率。

WO 2016/202308 A3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/094954

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F 17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, DWPI, VEN: template; KINGSOFT; JIAO, Guoqiang; regular, expression, match+, character string, file, name, recognize+, model, threshold

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 104881496 A (BEIJING KINGSOFT SECURITY SOFTWARE CO., LTD.), 02 September 2015 (02.09.2015), description, paragraphs [0110]-[0247], and figures 1-10	1-18
PX	CN 104881495 A (BEIJING KINGSOFT SECURITY SOFTWARE CO., LTD.), 02 September 2015 (02.09.2015), description, paragraphs [0098]-[0226], and figures 1-10	1-18
X	CN 104182531 A (GUANGZHOU KINGSOFT NETWORK TECHNOLOGY CO., LTD.), 03 December 2014 (03.12.2014), description, paragraphs [0037]-[0091] and [0100]-[0111], and figures 1 and 3	1-18
X	US 2013110585 A1 (INVISIBLEHAND SOFTWARE LTD.), 02 May 2013 (02.05.2013), description, paragraphs [0034]-[0192], and figures 1-12	1-18

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 November 2016 (11.11.2016)	**21 November 2016 (21.11.2016)**
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer **WANG, Kun** Telephone No.: (86-10) **62412187**

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/094954

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104881496 A	02 September 2015	None	
CN 104881495 A	02 September 2015	None	
CN 104182531 A	03 December 2014	None	
US 2013110585 A1	02 May 2013	None	

国际检索报告

国际申请号
PCT/CN2016/094954

A. 主题的分类

G06F 17/30(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F 17/30

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用）)

CNABS, CNTXT, CNKI, DWPI, VEN 金山，焦国强，正则，表达，匹配，字符串，文件，名，识别，模板，阈值，KINGSOFT, JIAO GUOQIANG, regular, expression, match+, character string, file, name, recogniz+, model, threshold

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 104881496 A (北京金山安全软件有限公司) 2015年 9月 2日 (2015-09-02) 说明书第[0110]-[0247]段，图1-10	1-18
PX	CN 104881495 A (北京金山安全软件有限公司) 2015年 9月 2日 (2015-09-02) 说明书第[0098]-[0226]段，图1-10	1-18
X	CN 104182531 A (广州金山网络科技有限公司) 2014年 12月 3日 (2014-12-03) 说明书第[0037]-[0091]、[0100]-[0111]段，图1、3	1-18
X	US 2013110585 A1 (INVISIBLEHAND SOFTWARE LTD) 2013年 5月 2日 (2013-05-02) 说明书第[0034]-[0192]段，图1-12	1-18

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2016年 11月 11日	2016年 11月 21日

ISA/CN的名称和邮寄地址	受权官员
中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088	王昆
传真号 (86-10)62019451	电话号码 (86-10)62412187

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2016/094954

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	104881496	A	2015年 9月 2日	无	
CN	104881495	A	2015年 9月 2日	无	
CN	104182531	A	2014年 12月 3日	无	
US	2013110585	A1	2013年 5月 2日	无	