



(12)发明专利申请

(10)申请公布号 CN 110770735 A

(43)申请公布日 2020.02.07

(21)申请号 201880039925.3

(74)专利代理机构 北京市金杜律师事务所
11256

(22)申请日 2018.05.21

代理人 黄倩

(30)优先权数据

15/621,009 2017.06.13 US

(51)Int.Cl.

G06F 40/103(2020.01)

(85)PCT国际申请进入国家阶段日

G06F 40/126(2020.01)

2019.12.13

G06F 40/131(2020.01)

(86)PCT国际申请的申请数据

PCT/US2018/033573 2018.05.21

G06F 40/221(2020.01)

(87)PCT国际申请的公布数据

WO2018/231427 EN 2018.12.20

(71)申请人 微软技术许可有限责任公司

地址 美国华盛顿州

(72)发明人 易晟 N·尼科利克 涂霄

I·斯托吉尔杰克维克

J·默加塞维克

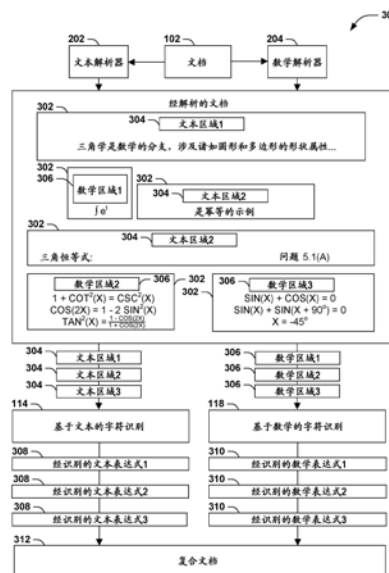
权利要求书2页 说明书17页 附图9页

(54)发明名称

具有嵌入式数学表达式的文档的编码转换

(57)摘要

符号识别技术可以被应用于包括各种形式的内容的文档。包括文本和数学表达式两者的文档可能会出现问題,因为应用与内容不匹配的识别器可能会产生异常结果。替代地,解析器可以基于每种类型的内容的特性来评估文档,以将各个区域分类为文本区域或数学区域中的一个区域。与每个区域的内容相对应的识别器可以被应用以产生包括经识别的文本表达式和经识别的数学表达式的复合文档。附加功能性可以基于经识别的内容被呈现;例如,诸如拼写检查和自然语言翻译的基于文本的工具被应用于包括文本表达式的文本区域,并且诸如数学方程式编辑器、数学表达式求解器和自动分配评级的基于数学的工具被应用于包括数学表达式的数学区域。



1. 一种评估文档的设备,所述设备包括:
处理器,以及
存储指令的存储器,所述指令在由所述处理器执行时,使所述设备:
将至少一个解析器应用于所述文档以标识所述文档中的数学表达式;
至少部分地基于所述数学表达式,将所述文档的区域分类为数学区域和文本区域中的一个区域;
将文本识别器应用于被分类为文本区域的所述区域,以生成至少一个经识别的文本表达式;
将数学识别器应用于被分类为数学区域的所述区域,以生成至少一个经识别的数学表达式;以及
编译包括所述至少一个经识别的文本表达式和所述至少一个经识别的数学表达式的复合文档。
2. 根据权利要求1所述的设备,其中标识所述文档中的所述数学表达式还包括:将由等式连接符连接的行内符号序列标识为数学表达式。
3. 根据权利要求1所述的设备,其中标识所述文档中的所述数学表达式还包括:
将主要包括数学符号的行内符号序列标识为数学表达式;以及
将以文字排列并由空格分隔的符号的序列标识为文本表达式。
4. 根据权利要求1所述的设备,其中标识所述文档中的所述数学表达式还包括:
将数学结构中的符号的序列标识为数学表达式;以及
将段落结构中的符号的序列标识为文本表达式。
5. 根据权利要求1所述的设备,其中标识所述文档中的所述数学表达式还包括:
跟踪用户的手写过程;
将在实质上线性序列中的一组手写符号标识为文本表达式;以及
将在实质上非线性序列中的一组手写符号标识为数学表达式。
6. 根据权利要求1所述的设备,其中标识所述文档中的所述数学表达式还包括:
将所选择的区域标识为数学区域和文本区域中的一个区域;以及
扩展所述所选择的区域以包括附加符号,所述附加符号位于所述所选择的区域附近并且类似于所述所选择的区域内的符号。
7. 根据权利要求1所述的设备,其中标识所述文档中的所述数学表达式还包括:
选择所述文档的符号的集合;以及
将符号的所述集合映射到数学表达式格式。
8. 根据权利要求1所述的设备,其中标识所述文档中的所述数学表达式还包括:
将第一区域标识为数学区域和文本区域中的一个区域;以及
将所述第一区域内的第二区域标识为与所述第一区域不同的区域类型。
9. 根据权利要求1所述的设备,其中对所述文档的所述区域进行分类还包括:仅在区域大于区域大小阈值的情况下,将所述区域分类为数学区域和文本区域中的一个区域。
10. 根据权利要求1所述的设备,其中对所述文档的区域进行分类还包括:
将区域的分类置信度标识为数学区域和文本区域中的一个区域;以及
仅在所述分类置信度大于分类置信度阈值的情况下对所述区域进行分类。

11. 一种评估文档的系统,所述系统包括:
文档解析器,所述文档解析器:
将至少一个解析器应用于所述文档以标识所述文档中的数学表达式,以及
至少部分地基于所述数学表达式,将所述文档的区域分类为数学区域和文本区域中的一个区域;
文档识别器,所述文档识别器:
将文本识别器应用于被分类为文本区域的所述区域,以生成至少一个经识别的文本表达式,以及
将数学识别器应用于被分类为数学区域的所述区域,以生成至少一个经识别的数学表达式;以及
文档编译器,所述文档编译器编译包括所述至少一个经识别的文本表达式和所述至少一个经识别的数学表达式的复合文档。
12. 根据权利要求11所述的系统,其中所述文档解析器通过以下方式将所述至少一个解析器应用于所述文档:
将文本解析器和数学解析器应用于所述文档;以及
从包括所述文本分析器和所述数学分析器中的一个解析器的所选择的解析器中接收至少一个区域,所述至少一个区域被所述所选择的分析器声称为要被分类为内容区域的区域,所述内容区域包括与所述所选择的分析器相匹配的内容。
13. 一种评估文档的方法,所述方法涉及具有处理器的设备,并且包括:
由所述处理器执行指令,所述指令使所述设备:
将至少一个解析器应用于所述文档以标识所述文档中的数学表达式;
至少部分地基于所述数学表达式,将所述文档的区域分类为数学区域和文本区域中的一个区域;
将文本识别器应用于被分类为文本区域的所述区域,以生成至少一个经识别的文本表达式;
将数学识别器应用于被分类为数学区域的所述区域,以生成至少一个经识别的数学表达式;以及
编译包括所述至少一个经识别的文本表达式和所述至少一个经识别的数学表达式的复合文档。
14. 根据权利要求13所述的方法,其中执行所述指令还使所述设备:
针对相应的数学区域,呈现针对数学工具的数学用户界面;以及
针对相应的文本区域,呈现针对文本工具的文本用户界面。
15. 根据权利要求14所述的方法,其中所述数学工具还包括:数学表达式求解器,所述数学表达式求解器在所述文档的所述至少一个数学区域中求解所述至少一个经识别的数学表达式。

具有嵌入式数学表达式的文档的编码转换

背景技术

[0001] 在计算的领域内,许多场景涉及包括多种内容的聚合的文档,多种内容诸如是键入文本、手写文本和绘图、数学方程式、嵌入图像、视频以及其他形式的媒体和数据。文本可以包括多种格式,诸如段落结构;被格式化为行和列集合的表格数据;以及边距和缩进。格式化也可以被应用于内容相对于彼此的位置集合;例如,可以应用行内布局,该行内布局使一些内容(诸如文本)提供容纳其他内容(诸如图像)的空间。

[0002] 在这样的场景中,识别技术可以被应用于标识文档的一部分的内容,并且将特定形式的内容翻译成具有更大结构程度的其他内容。作为第一个这样的示例,手写识别技术可以接受手写文本的图像,和/或可以检测来自诸如触笔的触摸设备的实时输入,并且可以将图像的笔触和/或输入翻译成已识别的字母数字符号集合,然后其可以以标准化的格式(诸如ASCII或Unicode)被编码。作为第二个这样的示例,可以对文档中的字母数字字符(包括以标准化格式编码的字母数字字符)进行语义评估,以标识表达式的语义内容,这可以使表达式翻译成不同的语言。可以利用许多这样的技术来识别文档的内容并采取适当的响应动作。

发明内容

[0003] 提供本发明内容以简化的形式介绍一些概念,这些概念将在下面的详细描述中进一步描述。本发明内容既不旨在标识所要求保护的的主题的关键因素或必要特征,也不旨在被用来限制所要求保护的的主题的范围。

[0004] 适用于文档的内容的识别技术常常利用特性集合,该特性集合描述其中所编码的表达式的类型。作为第一个这样的示例,用于英语表达式的自然语言识别技术可以基于此类书写的表达式的各种特性而进行开发,诸如利用拉丁字母的符号;段落格式样式,具有包括字符的序列的表达式,这些字符的序列包括用空格隔开并以从左到右的方式书写且以标点符号结尾的单词;以及公知单词的词典。识别器可以利用这些特性来指导文档中的英语表达式的识别;例如,可以通过将选项与已知单词的词典进行比较,并且优先选择词典中相对于不存在的单词的存在的单词,图像中使单词中的特定字符变得歧义的污迹可以被解决。作为第二个这样的示例,可以基于书写的数学表达式的诸如数字序列、数学符号和函数名称(例如,正弦函数)之类的各种特性、利用将符号分组为诸如方程式、分数和指数之类的数学结构的聚类来开发用于数学表达式的数学表达式识别技术。识别器可以利用这些特性来指导文档中的数学表达式的识别;例如,紧密接近的符号可以被识别为乘积,而间隔较远的符号可以被识别为不同的表达式。

[0005] 要被评估的文档可以利用识别器,该识别器适于文档中包含的内容的类型。但是,在一些情况下,内容的类型可能是未知的;例如,文档可能既包含自然语言文本又包含数学表达式。在一些情况下,自然语言文本和数学表达式可以利用相似的符号和格式;例如,字符序列“sin²”可以被识别为带有上标“2”字符(也许标示脚注或尾注)的英语单词“sin”,或者可以被识别为具有平方指数的Sine函数的数学表达式。将文本识别器应用于包括此字符

序列的整个文档可能导致纯基于文本的识别和输出,并且可能无法提供与基于数学的表达式相关的功能性,诸如提供数学表达式评估器。相反,如果将数学表达式识别器应用于将该字符序列识别为数学表达式的文档,则输出如果旨在作为文本,则可能无法正确解释该表达式,诸如将上标与脚注或尾注链接以提供书目条目。

[0006] 对序列的错误识别中可能还会出现其他问题;例如,如果包括积分符号 \int 的数学表达式被行内插入到文档中,则将符号与拉丁字母匹配的基于文本的识别器的应用可能会将符号标识为其最接近的字母数字字符,诸如大写S。结果,诸如“ $\int e^t$ ”之类的数学表达式可以被识别为英语单词“Set”,从而在已识别的文档中产生意想不到的且常常是无意义的输出。当将数学表达式识别器应用于还包括英语表达式的文档时,可能会出现另外的问题,诸如:“ $\int e^t$ is an example of idempotency”(“ $\int e^t$ ”是幂等的一个示例)。数学表达式识别器可能会开始评估数学表达式,但可能会尝试将以下文本标识为数学表达式的延续,从而产生无意义的识别结果。

[0007] 本文提出的是用于评估具有嵌入式数学表达式的文本文档的技术。根据这样的技术,文档解析器可以将至少一个解析器应用于文档以标识文档中的数学表达式,并且至少部分地基于数学表达式,将文档的区域分类为数学区域和文本区域中的一个区域。文档识别器可以将文本识别器应用于被分类为文本区域的区域,以生成至少一个经识别的文本表达式,并且可以将数学识别器应用于被分类为数学区域的区域,以生成至少一个已识别的数学表达式。文档编译器可以编译至少一个经识别的文本表达式和至少一个经识别的数学表达式,以生成包括经识别的文本区域和经识别的数学区域两者的复合文档。

[0008] 以这种方式对文档的处理可以使文档能够被识别为文本区域和数学区域的聚合。然后可以存储复合文档,该复合文档包括用于相应文本区域的经识别的文本表达式和用于相应数学区域的经识别的数学表达式。附加功能性可以源自于文本表达式和数学表达式的识别。例如,可以针对相应的文本表达式提供基于文本的工具,诸如基于文本的拼写检查和语法检查;实现基于文本的搜索的文本索引技术;总结文本表达式的内容的文本总结技术;以及将文本表达式从初始语言翻译成另一种语言的翻译技术。另一方面,可以针对相应的数学区域提供基于数学的工具,诸如帮助用户制定数学表达式的数学表达式编辑器;对数学表达式执行数学评估并计算结果的数学表达式求解器;验证数学表达式的正确性的数学表达式检查器,诸如自动家庭作业检查;以及将数学表达式可视化的数学可视化技术,诸如绘图或三维渲染。根据本文提出的技术,许多类型的功能性可以源自于文档中的文本表达式和数学表达式的识别。

[0009] 为了实现前述和相关目的,以下描述和附图阐述了某些说明性方面和实现。这些仅指示可以采用一个或多个方面的各种方式中的一些。当结合附图考虑时,根据以下详细描述,本公开的其他方面、优点和新颖特征将变得显而易见。

附图说明

[0010] 图1是以用于识别文档的内容的各种技术为特征的示例场景的图示。

[0011] 图2是根据本文提出的技术的、以用于解析文档的内容的各种技术为特征的示例场景的图示。

[0012] 图3是根据本文提出的技术的用于识别文档的内容的各种示例技术的图示。

[0013] 图4是根据本文提出的技术评估文档的内容的示例设备和系统的图示。

[0014] 图5是根据本文提出的技术的评估文档的内容的示例方法的图示。

[0015] 图6是根据本文提出的技术的使设备能够在虚拟环境内呈现应用的示例计算机可读存储设备的图示。

[0016] 图7是以如下示例为特征的示例场景的图示,这些示例是根据本文提出的技术来解析文档的内容的示例。

[0017] 图8是以如下示例为特征的示例场景的图示,这些示例是根据本文提出的技术来解析文档的内容的示例。

[0018] 图9是以如下示例为特征的示例场景的图示,这些示例是对已经根据本文提出的技术进行分类的各种区域的选择进行响应的自适应用户界面的示例。

具体实施方式

[0019] 现在参考附图描述所要求保护的主体,其中相似的参考标号始终被用来指代相似的元件。在下面的描述中,出于解释的目的,阐述了许多具体细节以便提供对所要求保护的主体透彻理解。然而,显然的是,可以在没有这些具体细节的情况下实践所要求保护的主体。在其他实例中,以框图的形式示出了结构和设备,以便促进描述所要求保护的主体。

[0020] A. 引言

[0021] 图1是示例场景100的图示,其特征在于文档102的呈现和评估。在该示例场景100中,文档102包括内容104的各个部分,这些部分呈现关于数学的一组表达式,其中内容104中的一些内容包括文本表达式106(例如,数学主题的英语自然语言描述),并且内容104中的一些内容包括数学表达式108(例如,公式和等式)。内容104可以被提供为键入文本和/或手写文本的图像,诸如教室白板的照片或数学教科书的扫描页面。内容104可以以各种方式(例如,作为段落、块、表格或图表)被布置,并且文档102可以包括其他类型的内容,诸如图像、视频和嵌入式数据文档。

[0022] 在此示例场景100中,可能期望解释文档102的内容104,以便提供一些附加功能性。作为第一示例,图像通常包括比编码符号(诸如ASCII或Unicode)大得多的数据量。如果大量文档102被提供(例如,学术图书馆的内容),则通过将文档102的内容104表示为编码符号而不是图像,可以大大降低存储要求。作为第二示例,可能期望能够对内容104进行索引和搜索,诸如在文档中找到术语“幂等”的所有实例,或者数学表达式“ $\sin^2(x)$ ”的所有实例。作为第三示例,可能期望为内容的各个类型104提供特定功能性,诸如应用于数学表达式108的数学表达式求解器,以及针对文本表达式106的拼写检查和/或语法检查工具。

[0023] 因此,可以利用识别技术来解释和提取文档102的内容104。可以利用多种识别技术来确定文档102和内容104的各种属性,诸如字体样式和大小、以及格式和布局。识别技术的应用可以导致对符号的序列的提取,该符号的序列对文档102的内容104进行编码。此类识别技术通常利用关于内容104的特性集合。例如,英语中的文本表达式106被特性地提供为拉丁字母中的一组字母数字符号,其包括从左到右和从上到下的序列中的单词。还可以利用格式化特性,诸如以段落格式特性地提供语句,其可以最初被缩进并且可以符合诸如左对齐、居中对齐或右对齐之类的对齐样式。数学表达式108经常展现出不同的特性集合;例如,以少量空白而被分组在一起的两个符号可以表示乘积,诸如表示两个三角函数的乘

积的“ $\sin(x) \cos(x)$ ”；而较大量的空白——诸如“ $\sin(x) \cos(x)$ ”仅表示不同且无关的函数。

[0024] 识别技术经常利用多种这样的特性来在其准确性中以高度的分类置信度来指导对可能表示内容104的符号的识别。这种识别的重大挑战是数学表达式108所利用的特性不仅可能不同于文本表达式106所利用的特性，而且实际上可能会冲突。作为第一个这样的示例，数学表达式108和文本表达式106通常利用不同组的符号。例如，希腊符号 μ 经常被利用在数学表达式中（例如，作为用于电磁渗透性的标准符号），但通常不会出现在英语文本表达式106中。如果根据英语文本表达式106的符号集合进行解释，该符号的实例可能被无意地解释为拉丁字母小写的U符号。相反，英语文本表达式106有时包括诸如段落符号(¶)和版权符号(©)的符号，但是这些符号通常不出现在数学表达式108中，并且可能被解释为不同的数学符号。作为第二个这样的示例，尽管英语文本表达式106经常被分组到一个段落中，其中在一行上的第一句的结尾可以紧随第二句的开头，但是数学表达式108很少以这样的方式进行组织。相反，数学表达式108利用诸如分数的格式化结构，其包括在同一行中竖直对齐但由分数线分隔的分子和分母，而英语文本表达式106不具有等效的格式化结构。作为第三个这样的示例，一些特性可以在数学表达式108和文本表达式106两者中被利用，但是可能表示不同的语义。例如，数学表达式108中的诸如“ \sin^2 ”的表达式可以标示具有平方指数的三角函数，但是文本表达式106中的相同表达式可以标示英语单词“sin”和将单词与脚注或尾注相关联的参考数字。

[0025] 鉴于这些区别，可能期望针对文本表达式106利用基于文本的字符识别技术112并且针对数学表达式108利用基于数学的识别技术114。然而，在诸如在图1的示例场景100中所呈现的文档102中可能会使该任务复杂化。

[0026] 将基于文本的字符识别器112应用于文档102可能会在所产生的基于文本的识别114中导致许多异常。作为第一个这样的示例，诸如“ f_e ”之类的数学表达式108可能会被映射到最接近的匹配英语符号以产生英语单词“Set”。基于文本的字符识别技术112可能将此标识为有效翻译，这是因为它与已知的英语单词相匹配，并且因为数学表达式108是与文本表达式106一起被顺序地呈现。结果，基于文本的识别114可能包括错误翻译124，其对于读者而言可能难以理解。作为第二个这样的示例，应用已知术语的词典可能会导致一些正确翻译的符号被标记为印刷错误126，诸如通常被用来指示数学表达式108中的余割函数的缩写“csc”，但是其在英语106中不是被接受的单词。作为第三个这样的示例，数学表达式108中的指数120的各种实例可能会被翻译为格式化，诸如在文本翻译中将指数120放置在单独的一行128上，或者被翻译为标示脚注而不是指数的上标。作为第四个这样的示例，一些符号可能会简单地被标识为错误130；例如，数学表达式108可以以全高符号开头，但是然后包括具有由分数线分开半角符号的分数122，但基于文本的字符识别技术112可能未被配置为预期或正确评估分数，因为这种符号构造在英语中不存在。替代地，基于文本的识别技术112可能会尝试将分数122映射到英语符号，并且以任何合理的分类置信度都无法映射，可能在基于文本的识别114中的该点处指示错误。

[0027] 类似地，将基于数学的字符识别器116应用于文档102可能会在所产生的基于数学的识别118中导致许多异常。作为第一个这样的示例，数学表达式108“ f_e^t ”可以被正确地翻译，但是继续尝试将随后的符号映射到数学符号可能会导致错误翻译124（例如，将单词

“is”映射到数字“15”)，并可能最终导致错误130，其中利用可接受的分类置信度无法实现这种映射。作为第二个这样的示例，数学格式化语义的应用可能导致一些非数学符号的错误翻译124。例如，短语“Problem 5.1 (A)” (问题5.1 (A)) 在英语中可能只是提供对家庭作业问题的列举，但是如果根据数学表达式108的格式化特性来解释，则该短语可以指示值(A)与数值5.1的乘积。这些和其他形式的错误翻译124可能来自于对文档102的基于文本的识别技术112和基于数学的识别技术116的“天真”应用。

[0028] B. 提出的技术

[0029] 图2-图3一起是示例场景200的图示，其特征在于根据本文提出的技术的对图1的文档102的解释。

[0030] 如图2的示例场景200中所图示，文档102由文本解析器202评估以标识文档102中的文本表达式106，并且由数学解析器204评估以标识文档102中的数学表达式108。例如，文档102中的内容104的各种集合可以被提供给解析器，并且每个解析器可以指示内容104可能包括该解析器识别的表达式类型的分类置信度206。例如，包括第一组三角学主题的叙述性描述的内容104可以由解析器评估，并且文本解析器202可以指示高文本表达式分类置信度206，其指示内容104非常可能是文本表达式106，而数学分析器204可以指示低数学表达式分类置信度206，其指示内容104不太可能是数学表达式108。这种确定可以涉及例如：对内容104的布局和/或格式化是否是文本表达式106和/或数学表达式108的特性的初步评估；一些符号的标识和频率，这些符号高度具有文本表达式106和/或数学表达式108的特性（例如，将单词“the”的实例作为英语文本表达式106的强特性进行查找，并将符号“=”的实例作为数学表达式108的强特性进行查找）；和/或内容104与先前已被标识为文本表达式106或数学表达式108的其他内容104的接近度和/或相似性。基于由文本解析器202和数学解析器204报告的分类置信度206，可以将内容104标识为可能包括文本表达式106。以类似方式处理文档102的其他内容104。例如，被分组在一起的一组内容106（诸如等式的序列）可以被各个地标识为数学表达式108（根据数学解析器204报告的高分类置信度206，以及文本解析器202报告的低分类置信度206）。

[0031] 如图2的示例场景200中所图示，以这种方式对文档102进行解析可能会产生一些实例。作为第一个这样的示例，解析器可能会将一些内容104标识为包括：可能是数学表达式108（例如“ je^t ”）的第一部分，数学解析器204报告其为高分类置信度206并且文本解析器202报告其为低分类置信度206；和可能是文本表达式106（例如，“is an example of idempotency”（是幂等的一个示例））的第二部分，数学解析器204报告其为低分类置信度206并且文本解析器202报告其为高分类置信度206——尽管有不寻常的术语“幂等”。因此，可以将内容104标识为包括根据解析器而被不同地标识的不同部分。相反，在图2的示例场景200中图示出的解析技术可以促进对潜在的歧义表达式的准确确定。例如，短语“Problem 5.1 (A)” (问题5.1 (A)) 可以被文本解析器202标识为仅具有中等分类置信度206的文本表达式106（因为符号“5.1 (A)”不是典型的英语文本表达式106），但是可以被数学解析器204标识为具有甚至更低分类置信度206的数学表达式108（因为单词“Problem”（问题）不在一组可标识的数学符号之中）。因此，该表达式可以被正确地标识为更可能是文本表达式106而不是数学表达式108。因此，以比较的方式对文档102的内容104进行解析可以使得能够标识其中提供的文本表达式106和数学表达式108。

[0032] 图3呈现了示例场景300的图示,其特征在于使用图2的比较解析示例来完成对文档102的评估。文本解析器202和数学解析器204应用于文档102的内容104使得能够标识文本表达式106和数学表达式108。这种标识可以使得能够将解析的文档的各个区域302分类为包括文本表达式106的文本区域304或者包括数学表达式108的数学区域306。例如,各个区域302可以被标识为相似标识的表达式的近似集合,诸如英语文本表达式106的段落或数学表达式108的块。在一些实例中,特定区间可以被标识为包括两个不同分类的区域302,诸如数学区域306,然后是文本区域304(例如,“ $f e^t$ ”然后是“is an example of idempotency”(是幂等的一个示例)),或者用作数学区域306标题的文本区域304(例如,“Problem 5.1 (A)” (问题5.1 (A)))。在一些实例中,分类可以导致在封闭的文本区域304内被标识的数学区域306,或者在封闭的数学区域306内被标识的文本区域304。在其他实例中,区域302的基于接近度的标识还可以促进对其中分类的分类置信度206不清楚的区域302的消歧;例如,在文本表达式206的一组高分类置信度分类内,可能被歧义地标识为文本表达式206或数学表达式208的一个短语(诸如短语“ \sin^2 ”)可能会作为文本区域304的一部分而偏向于标识为文本表达式206。

[0033] 以这种方式标识的解析文档的区域302可以使用对应的识别器而被识别。例如,文本区域304可以被提供给基于文本的识别技术114,并且数学区域306可以被提供给基于数学的识别技术118,以便产生经识别的文本表达式308和经识别的数学表达式310,其分别以符号格式被,该符号格式编码适于内容的每种类型。然后,根据本文提出的技术,可以将经识别的文本表达式308和经识别的数学表达式310编译成复合文档312,该复合文档312包括在原始文档102的内容104中呈现的文本表达式106和数学表达式108两者。

[0034] C. 技术效果

[0035] 在文档评估领域中使用本文提出的技术可以提供多种技术效果。

[0036] 通过使用本文提出的技术可以实现的第一技术效果涉及:在对包括文本表达式106和数学表达式108两者的文档102的内容104的识别中的更高层次的准确性。如在本文中尤其是在图1的示例场景100中所指出的,仅通过应用基于文本的识别技术112或基于数学的识别技术114来识别文档102可能会导致许多异常——尤其是对于可能被合理识别为一种类型的表达式、但是能被更准确地识别为另一种类型的表达式的那些表达式而言(例如,错误地将数学表达式108“ $f e^t$ ”识别为包括单词“Set”的文本表达式106)。也就是说,应用基于文本的识别技术112可能会产生较低分类置信度,但是仍然可以产生将该表达式作为文本表达式106的合理识别,而应用数学解析器204可以将这组符号正确标识为数学表达式108。一个或多个解析器的应用还可以标识混合表达式的实例,诸如数学表达式108后接文本表达式106的行内组合,并且将区域302分别分类为数学区域306和文本区域304可以使得能够应用与相应区域302的内容104相匹配的识别器。如在图1和图2-图3的比较示例中一般图示出的,对文档102应用一个或多个解析器,以标识文本区域304和数学区域306可以促进对文档102的内容104的准确识别。

[0037] 可以通过使用本文提出的技术来实现的第二技术效果涉及:基于解析的结果对呈现给用户的文档编辑系统的用户界面进行适配。一些文档编辑环境可以提供协助用户输入文本表达式106的第一编辑工具集合(例如,拼写检查和语法检查工具)和协助用户输入数学表达式108的第二编辑工具集合(例如,表达式编辑器,其允许用户限定格式化结构,诸如

括号、指数和分数)。然而,许多这样的文档编辑环境要求用户肯定地指示用户当前正在输入的内容104的类型,和/或选择用户希望使用的编辑工具集合。如果用户经常在内容的类型之间进行切换,并且因此每次区域转换时都必须更改由文档编辑环境提供的工具集合,则这些要求可能会变得很麻烦。根据本文中呈现的技术对文档102的内容104的解析可以使得能够以动态方式自动选择呈现给用户的工具集。当用户开始在新区域302中输入内容104或选择现有区域302以插入或编辑内容104时,可以将解析应用于内容104,以确定区域302是包括文本区域304还是数学区域306,并且文档编辑环境可以针对所标识的内容104自动地选择并呈现对应的编辑工具集合。当第一类型的区域302中的内容104被更改为另一类型的内容104时(例如,当用户最初正在文本区域304中输入文本表达式106,然后切换到输入数学表达式108时,提示标识新的数学区域306),解析也可以实现自动标识。作为另一个这样的示例,可以通过提供区分经识别的符号和未识别的符号的拼写、句法、和语法检查服务来促进内容104的输入;例如,可以将英语术语词典应用于文本区域304中的单词拼写,而可以利用数学术语词典来验证数学区域306中的单词拼写。以这种方式,本文提出的解析技术可以实现或促进用户界面的适配,其帮助用户将内容104输入到文档102中。

[0038] 通过使用本文提出的技术可以实现的第三技术效果涉及:基于在每个区域302中提供的内容104的自动标识,来呈现附加的特定于内容的工具和功能性。例如,经识别的文本表达式308可以由文本到语音处理器、翻译服务以及汇总服务进一步处理,该文本到语音处理器读取文本表达式,以便促进可访问性;该翻译服务将文本表达式从母语翻译成另一种语言,诸如自动将英语文本表达式106翻译成西班牙语文本表达式106;而该汇总服务自动评估文本表达式106的语义内容并准备摘要。此外,经识别的数学表达式310可以由数学方程式求解器、数学表达式检查器以及数学可视化器进一步处理,该数学方程式求解器计算并呈现数学表达式108的解;该数学表达式检查器验证输入的数学表达式108的准确性,诸如自动数学家作业评级服务;而该数学可视化器准备并呈现数学表达式108的可视化,诸如函数图和方程组的二维和三维模型。以这种方式,除了协助用户输入和编辑各种类型的表达式之外,本文提出的技术还可以使得各种补充服务能够根据本文提出的技术为文档102的内容104提供上下文相关的功能性。

[0039] 通过本文提出的技术可以实现的第四技术效果涉及:文档评估过程的性能。如图2-图3的示例实施例中所呈现的识别过程的组织可以促进文档编辑场景的许多方面,特别是包括本文提出的其他技术效果,在图2-图3中将文档102的解析与对文档102的各个区域302的内容104的识别区分开。而对文档102的识别被执行,以将相应区域302的整体内容104翻译为以标准化格式编码的一组符号(例如,将文本区域304的整个文本表达式106翻译为ASCII或Unicode),提供文档102的解析只是为了将相应区域302标识为文本区域304或数学区域306。在比较这些任务时,可以理解的是,与内容104的识别相比,解析可以包括相对简单的任务,并且解析的结果在较短的时间范围内是可实现的,这可以使用户对数据的输入与至少部分地基于此确定的功能性(诸如与用户当前正在创建和/或编辑的内容104相匹配的用户界面的呈现)之间的延迟降低。以这种方式,将文档评估过程组织为不同的解析和识别任务,可以使得能够基于对应结果的效用以不同的方式来实现这些任务的应用。例如,可以基本上连续和实时的方式来应用解析过程(例如,频繁地评估和/或重新评估由用户创建和/或编辑的内容104,以将区域302分类为文本区域304和数学区域306中的一个区域),

以实现高度响应的用户界面,而可以在延迟和/或脱机的基础上提供内容104的识别,以将文档102翻译为标准化的符号编码。根据本文提出的技术可以设想出许多这样的技术效果,这些技术效果可以出自于文档标识过程的组织。

[0040] D. 示例实施例

[0041] 图4是示例场景400的图示,其特征在于根据本文提出的技术评估文档102的内容104的一组示例实施例。

[0042] 在该示例场景400中,提供了示例设备402,其包括处理器404和存储文档102的存储器406,以及评估文档102的内容104的示例文档评估系统408。可以由用户416例如使用键盘、诸如触笔的手写笔工具或解释用户416的口头表达的语音处理器来创建文档102的内容104。备选地或附加地,文档102的内容104可以由其他人创建并由用户416捕获,例如使用设备402的照相机来捕获教室白板或教科书页面的视觉内容104。图4的示例场景400中的示例文档评估系统408包括存储在存储器406中的指令,该指令在由处理器404执行时使示例设备402应用本文提出的技术来评估文档102的内容104,并生成经由示例设备402的显示器420呈现给用户416的复合文档312。此类实施例中的变化还可以使得能够使用本文提出的技术,诸如将所提出的技术的功能性重构到不同的软件架构中,和/或以不同的方式实现示例文档评估系统408的一个或多个元件,诸如分立电路或现场可编程门阵列(FPGA),其中此类备选文档评估系统也导致根据本文提出的技术对文档102进行评估。

[0043] 示例文档评估系统408包括文档解析器401,其将至少一个解析器应用于文档102以标识文档102中的数学表达式108。在该特定示例场景400中,文档解析器410将如下两者应用于文档102:标识文档102的内容104中的文本表达式106的文本解析器202,以及标识文档102的内容104中的数学表达式108的数学解析器204。至少部分地基于数学表达式108的标识(以及可选地,文本表达式106的附加标识),文档解析器410将文档102的区域302分类为文本区域304和数学区域306中的一个区域。示例文档评估系统408还包括文档识别器412,其将文本识别器114应用于被分类为文本区域304的区域302,以生成经识别的文本表达式308,并且将数学识别器118应用于被分类为数学区域118的区域302,以生成经识别的数学表达式310。示例文档评估系统408还包括文档编译器414,其编译包括经识别的文本表达式308和经识别的数学表达式310的复合文档312。以这种方式,示例文档评估系统408使得示例设备402能够根据本文提出的技术来评估文档。

[0044] 图5是本文提出的技术的第三示例实施例的图示,其被图示为评估文档102的内容104的示例方法500。示例方法500涉及具有处理器404的设备,并且可以例如被实现为存储在诸如固件、系统存储器、硬盘驱动器、固态存储组件、或磁性或光学介质之类的存储器406中的一组指令,其中,处理器404对指令的执行使该设备根据本文提出的技术进行操作。

[0045] 示例方法500从502开始,并且涉及在处理器404上执行504指令。特别地,执行指令使设备将至少一个解析器应用506于文档102,以标识文档102中的数学表达式108。执行指令还使设备至少部分地基于数学表达式,将文档102的区域分类508为数学区域306和文本区域304中的一个区域。执行指令还使设备将文本识别器114应用510于被分类为文本区域304的区域302,以生成至少一个经识别的文本表达308。执行指令还使设备将数学识别器118应用512于被分类为数学区域306的区域302,以生成至少一个经识别数学表达式310。执行指令还使设备编译514包括至少一个经识别的文本表达式308和至少一个经识别的数学

表达式310的复合文档312。以这种方式,示例方法500使设备能够根据本文提出的技术来评估文档102的内容104,并且因此在516结束。

[0046] 另一个实施例涉及一种计算机可读介质,该计算机可读介质包括被配置为应用本文提出的技术的处理器可执行指令。此类计算机可读介质可以包括各种类型的通信介质,诸如可以通过各种物理现象(例如,电磁信号、声波信号或光信号)以及在各种有线场景(例如,经由以太网或光纤电缆)和/或无线场景(例如,诸如WiFi之类的无线局域网(WLAN),诸如蓝牙之类的个人局域网(PAN),或蜂窝或无线网络)中传播并且对一组计算机可读指令进行编码的信号,这些指令在由设备的处理器执行时使该设备实现本文提出的技术。这样的计算机可读介质还可以包括(作为不包括通信介质的一类技术)计算机-计算机可读存储器设备,诸如存储器半导体(例如,利用静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)和/或同步动态随机存取存储器(SDRAM)技术的半导体)、一盘硬盘驱动器、闪存设备或磁盘或光盘(诸如CD-R、DVD-R或软盘)、该计算机-计算机可读存储器设备对一组计算机可读指令进行编码,这些指令在由设备的处理器执行时使该设备实现本文提出的技术。

[0047] 在图6中图示出了可以以这些方式设想的示例计算机可读介质,其中,实现600包括计算机可读存储器设备602(例如CD-R、DVD-R或一盘硬盘驱动器),在该计算机可读存储器设备602上对计算机可读数据604进行编码。该计算机可读数据604继而包括一组计算机指令606,这些指令在设备的处理器404上被执行时,使设备根据以下阐述的原理进行操作。作为第一个这样的示例,处理器可执行指令606可以提供评估文档102的内容104的系统,诸如图4的示例场景400中的示例文档评估系统408。作为第二个这样的示例,处理器可执行指令606的执行可以使设备执行评估文档102的内容104的方法,诸如图5的示例方法500。本领域普通技术人员可以设想许多这样的计算机可读介质,其被配置为根据本文提出的技术进行操作。

[0048] E. 变型

[0049] 本文讨论的技术可以在许多方面设想变型,并且一些变型相对于这些技术和其他技术的其他变型可以呈现附加的优点和/或减少缺点。此外,可以组合地实现一些变型,并且一些组合可以通过协同合作而具有附加的优点和/或减少的缺点。可以将变型并入到各种实施例中(例如,图4的示例设备402;图4的示例文档评估系统408;图5的示例方法500;和/或图6的示例计算机可读存储器设备600)以在这些实施例上赋予单个和/或协同优势。

[0050] E1. 场景

[0051] 在这些技术的实施例之中可以变化的第一方面涉及其中可以利用这种技术的场景。

[0052] 作为该第一方面的第一变型,本文提出的技术可以在多种设备上被利用,诸如服务器、工作站、膝上型计算机、控制台、平板电脑、电话、便携式媒体和/或游戏机、嵌入式系统、家用电器、车辆和可穿戴设备。此类设备还可以包括设备的集合,诸如分布式服务器场,该分布式服务器场可能在地理上分布的区域中提供多个服务器,该多个服务器可以互操作以代表诸如管理员、访客、顾客、客户和其他应用和/或设备指令的各种用户102来评估文档的内容104。另外,实现本文提出的技术的设备还可以包括各种输入设备,诸如键盘;诸如触笔或触敏显示器的手写式输入设备;接收语音输入的麦克风和语音处理器;以及接收文档102的图像的照相机或扫描仪,诸如教科书的页面或教室白板的照片。

[0053] 作为该第一方面的第二变型,本文提出的技术可以与各种文档102一起使用,诸如教科书、论文、白板插图、写入或键入的笔记、考试或家庭作业以及参考指南。文档102还可以包括从网络检索的数据,诸如网站的一个或多个网页的内容104。文档102可以限于文本表达式106和数学表达式108,或者还可以包括其他形式的内容104,诸如图像或嵌入式数据文档。文档102可以由设备的用户416创建,和/或由第三方创建并由用户416捕获和/或接收。文档102可以最初在实现本文提出的技术的设备上被创建,和/或可以在其他位置被创建然后由设备接收,诸如对教科书的扫描。作为一个示例,可以通过互联网提供文档评估服务,诸如基于云的文档分析服务,该文档评估服务评估通过诸如互联网的网络传输到该服务的文档102的内容104。作为另一示例,用户可以根据本文提出的技术利用第一设备来创建文档102并且利用第二设备来评估文档102。

[0054] 作为该第一方面的第三变型,通过本文呈现的技术处理的文档102可以提供各种内容104。作为第一个这样的示例,文本表达式106可以包括例如用诸如英语或西班牙语的自然语言最初叙述和/或书写的语句,和/或从其他来源翻译成自然语言的语句,诸如最初以初始语言提供然后翻译成自然语言的文本表达式106。文本表达式106也可以是不同语言的文本表达式106的聚合,诸如包括英语和西班牙语二者的文本表达式106的文档102。备选地或附加地,文本表达式106可以以其他类型的语言来呈现,诸如辅助语言或构造语言(例如,埃丝博兰多语或克林贡语);非语言的语言(例如手语或盲文);和/或功能语言(例如,编程语言)。文本表达式106可以由诸如设备的用户416的个体来创建,和/或可以由诸如散文生成算法的设备来创建。作为第二个这样的示例,数学表达式108还可以包括各种表达式,诸如由设备生成的语言数学表达式108(例如,口语方程式)和数学表达式108(例如,由计算机生成的代数相关方程组)。作为第三个这样的示例,文本表达式106可以最初与数学表达式108集成在一起,或者可以已经从一个或多个源(例如,补充有数学表达式108的原始文本文档)被聚合。

[0055] 作为该第一方面的第四变型,可以在体现本文提供的技术的设备中利用各种架构。这样的设备可以利用各种技术,诸如存储指令的易失性或非易失性存储器,当指令被处理器404执行时,该指令使计算设备利用本文提出的技术;实现本文提出的技术的功能性的一个或多个部分的一组分立电路组件;和/或混合解决方案,诸如现场可编程门阵列(FPGA),其合成基于功能指定的程序的结构电路。另外,实现本文提出的技术的设备、系统和方法可以利用对多个元素的各种组织,例如,这种方法的步骤的数目、顺序和/或层次关系可以变化,并且这种设备或系统的元件的数目、功能性和/或层次关系可以变化。作为第一个这样的示例,体现当前提出的技术的系统可以利用不同的元件,这些不同的元件包括文本解析器202、数学解析器204、文本识别器114和数学识别器118;其他系统可以组合一个或多个元素,诸如提供经组合的数学解析器204/数学识别器118,或经组合的文本解析器202/数学解析器204。作为第二个这样的示例,此类系统的组件可以按顺序操作(例如文本解析器202可以首先被应用于文档102,并且在其评估完成之后,可以应用数学解析器204)和/或可以同时操作(例如,文本解析器202和数学解析器204可以同时被应用于文档102)。其他实施例可以完全省略文本解析器202,并且可以假定未被数学解析器204声明的所有区域302默认都是文本表达式106。尽管如此,这种架构变型可以体现所提出的权利要求的元素。场景和架构中的所有此类变型都在本文提出的公开技术的范围内。

[0056] E2. 解析技术

[0057] 在本文提出的技术的实施例之中可以变化的第二方面涉及如下方式：将文档102的内容104解析为区域302的方式、将表达式标识为文本表达式106和数学表达式108中的一个表达式的方式、以及相应地各个区域302将分类到文本区域304和数学区域306中的一个区域的方式。多种因素可以被利用来通知每个这样的任务的确定。

[0058] 作为该第二方面的第一变型，解析可以涉及通过标识表达式中的一个或多个特性符号来将表达式标识为文本表达式106和数学表达式108中的一个表达式。作为第一个这样的示例，如果表达式包括作为文本表达式106的特性（并且其不是数学表达式108的特性）的符号，诸如像引号的标点符号，则可以将该表达式标识为文本表达式106。例如，如果表达式包括作为数学表达式108的特性（并且其不是文本表达式106的特性）的符号，诸如算术运算符（例如，符号“+”、“-”、“*”和/或“/”），包括由等式连接符（例如“<”、“>”或“=”）连接的行内符号序列，则可以将该表达式标识为数学表达式108。尽管该技术可能类似于传统的符号识别，但是可以认识到，该标识过程可能比完整的识别过程简单得多；此解析技术仅涉及标识是否存在较小的一组符号实例，而不是尝试将表达式的每个符号映射到ASCII或Unicode之类编码的已知符号，该较小的一组符号实例是可诊断的以用于将标识表达式为文本表达式106或数学表达式108。作为第二个这样的示例，解析可以涉及通过通常考虑在其中使用的符号的类型来将表达式标识为文本表达式106和数学表达式108中的一个表达式。例如，可以将主要由数字和算术运算符组成的表达式标识为数学表达式108，而可以将主要由被分成用空格分隔的单词的字母和标点符号组成的表达式标识为文本表达式106。尽管此解析技术可能涉及与已知符号的更高程度的匹配输入，但与完全识别技术相比，该技术可能仍然更加精简，完全识别技术还是寻求将符号集合映射到诸如英语中的已知单词之类的实体中。通过仅查看表达式中使用的符号的类型而不必考虑将符号集合翻译成的实体（诸如单词或数字），可以在更快、更有效的基础上执行此解析任务。

[0059] 作为该第二方面的第二变型，可以根据符号的格式、间隔和布局来实现表达式的解析。例如，数学表达式108通常被布置为主要包括数学符号的行内符号序列，而文本表达式106可以包括被布置为单词并且由空格分隔的符号序列。备选地或附加地，该解析可以基于以下确定：符号以作为提供在段落样式布局中的单词序列而被聚类，或是作为包括一个或多个数学结构（例如，指数和分数）的方程样式布局中的符号序列的特性方式而被聚类。这样的解析可以能够在不考虑这样格式化的实际符号的情况下区分文本表达式106和数学表达式108。

[0060] 作为该第二方面的第三变型，解析和/或识别可以基于诸如图像之类的符号的静态布置。备选地，解析和/或识别可以基于创建文档102的个体的输入字符或笔划的序列。例如，解析可以涉及跟踪用户的手写过程，并将在实质上线性序列中的一组手写符号（例如，英语手写序列通常遵循语言的从左到右和从上到下的布局）标识为文本表达式106。相反，解析可以通过跟踪用户的手写过程，将在实质上非线性序列中的一组手写符号标识为数学表达式108（例如，当书写诸如分数的数学结构并执行诸如加法、乘法和长除法之类的数学过程时，与公式化文本表达式106时相比，用户的书写可能展现出基本上更多的竖直变动幅度和向左移动）。

[0061] 作为此第二方面的第四变型，解析可以考虑将非符号与文本表达式106或数学表

达式108相关,该非符号与表达式相关联。例如,数学表达式108可以与数学绘图(诸如函数图)相关联。解析可以能够评估与表达式相关联的非符号内容,并且通过将内容标识为数学绘图,可以通过关联将表达式标识为数学表达式。尽管此类信息可能不是孤立处置的,但分析可能会全盘考虑此信息以及其他因素,以完成对文本表达式106或数学表达式108的确定。

[0062] 图7是示例场景700的图示,其特征在于根据本文提出的技术来解析文档102的各种表达式的内容104的示例。在该示例场景700中,提供了根据一组特性来评估表达式的文本解析器202,该一组特性诸如是其中经常提供文本表达式106的独特格式702(例如,段落样式)以及与数学表达式108相比更可能出现在文本表达式106中的一组符号704(例如,引号、逗号和问号)。类似地,提供了根据一组特性来评估表达式的数学解析器204,该一组特性诸如是其中经常提供数学表达式108的独特格式702(例如,行内方程式和包括分数的竖直对齐的表达式)以及与文本表达式106相比更可能出现在数学表达式108中的一组符号704(例如,算术运算符和等号)。

[0063] 可以按以下方式将解析器应用于各个区域302内的表达式。作为第一示例710,第一区域302可以包括一组表达式106,该一组表达式106体现段落格式702并且利用代表文本表达式106的符号704的若干实例。因此,文本解析器202可以提供表达式包括文本表达式106的高分类置信度206,而数学解析器204提供表达式包括数学表达式108的低分类置信度206。因此,该表达式可以被标识为文本表达式106。类似地,作为第二示例712,第二区域302可以包括一组表达式106,该一组表达式106体现行内格式702和小数122,并利用代表数学表达式108的符号704的若干实例。因此,文本解析器202可以提供表达式包括文本表达式106的低分类置信度206,而数学解析器204提供表达式包括数学表达式108的高分类置信度206。因此,该表达式可以被标识为数学表达式108。

[0064] 作为该第二方面的第五变型,解析可以使用表达式的接近度来辅助其他表达式的标识。作为第一个这样的示例,解析可以包括首先将所选择的区域302标识为数学区域306和文本区域304中的一个区域(例如,标识一组符号,该一组符号展现出包括诸如分数的数学结构或诸如方程式的数学表达式108的高分类置信度206),然后扩展所选择的区域302,以包括靠近所选区域302且与所选区域302内的符号相类似的其他符号。例如,符号序列“ \sin^2 ”如果单独被评估,则可能会被认为是歧义的,但是该符号序列“ \sin^2 ”与被标识为具有高分类置信度206的数学区域306的区域302的邻近度可以提高分类置信度206,其中该符号序列是作为数学区域306一部分的数学表达式108。相反,该符号序列与文本区域304的接近度可以提高分类置信度206,其中该符号序列是作为文本区域304的一部分的文本表达式106。因此,该解析可以通过扩展先前标识的区域302以包括其他附近的符号而被执行。这样的解析可以努力将相似内容的群组标识到文档102的各个区域中,诸如包括一组相似特性化符号的矩形。作为第二个这样的示例,解析可以涉及将第一区域302标识为数学区域306和文本区域304中的一个区域,并且将第一区域302内的第二区域302标识为与第一区域302不同的区域类型,诸如,特定类型的区域302被嵌入在不同类型的区域302内。例如,对爱因斯坦作品的叙述性描述(其在文本区域304内被标识为文本表达式106)可以包括对广义相对论的引用,如方程式“ $e=mc^2$ ”所表述的,其在数学区域306内被标识为数学表达式108。备选地或附加地,解析可以涉及仅对区域304大于区域阈值大小的区域进行分类。例如,如果

将诸如“a+b”之类的短表达式嵌入到相当长的文本表达式106中,则可能不希望将诸如“a+b”之类的短表达式分类为数学表达式108,其中将该表达式表征为文本表达式106的分类置信度206基本上不低于将表达式表征为数学表达式108的分类置信度206。

[0065] 图8是示例场景800的图示,其涉及将区域302分类为文本区域304和数学区域306中的一个区域。在该示例场景800中,在第一时间804处,文档102可以被评估以标识表示数学表达式108的符号708,诸如相等符号。

[0066] 在第二时间806处,可以将与经标识的相等符号708在水平行内的、并在经标识的相等符号708附近的符号与行内符号分组在一起,其可以被共同标识为数学表达式108。具有文本表达式106的特性的其他符号106可以类似地分组在一起并被标识为文本表达式106。附加地,可以将邻近符号802的歧义集合标识为是歧义的,即,包括具有文本表达式106和数学表达式108的特性的符号708。因此,可以以更高程度的粒度来评估该符号708的集合,从而导致将第一子组的符号708确定为数学表达式108并且将第二子组的符号708确定为文本表达式106。

[0067] 在第三时间808处,包围一组相似标识的表达式的矩形区域可以被限定,并且可以根据表达式而被分类。文本区域304可以被标识为文本表达式106的集合,并且数学区域306可以被标识为数学表达式108的集合。邻近符号802的歧义集合可以被划分为两个区域302,其中第一区域302包括数学区域306并且第二区域302包括文本区域304。以这种方式,可以以逐步的方式解析文档102的内容104,以对文档102内的所有符号和表达式进行分类。可以利用许多这样的技术来根据本文提出的技术执行对文档102的内容104的解析和对区域302的分类。

[0068] E3. 解析过程

[0069] 在本文提出的技术的实施例之中可以变化的第三方面涉及计算过程的组织,以便将本文描述的解析技术应用于文档102的内容104。也即是说,虽然可以针对实施例选择特定集合的解析技术,但是这种解析技术被实现为文档评估系统408的方式可以显著影响文档102的评估结果。

[0070] 作为该第三方面的第一变型,可以以各种方式将解析应用于文档102,尤其是在用户正在创建内容104时。例如,在一些变型中,例如作为连续的实时过程的一部分,内容104(以及可选的输入字符和/或笔划的序列)可以在内容104的创建期间被解析和/或识别。在其他变型中,可以将文档102的评估划分为内容捕获阶段(在该内容捕获阶段期间仅接收和存储内容104以及可选的内容创建的序列)和内容解析和识别阶段(在该内容解析和识别阶段期间,内容104以及可选的内容创建序列被分类为各种类型的区域302,然后使用对应的识别器进行识别)。内容解析和识别阶段可以在周期性的基础上发生(例如,每分钟发起一次解析和标识,以评估一批新提供的内容104);在空闲的基础上发生(例如,等待处理能力充足的空闲时刻,诸如500毫秒的空闲时间,然后对文档102的内容104进行内容解析和识别);和/或根据需要发生(例如,等待来自用户102的对解析和/或识别文档102的内容104的请求,和/或直到对包括文档102的用户的数据集调用了索引过程)。还在其他变型中,对文档102的评估可以被分配;例如文档102的评估可以被分配。例如,可以以近似连续的方式执行将区域302解析为文本区域304和数学区域306中的一个区域,而作为潜在的计算量更大的过程,则可能在延期的基础上发生对文档102的内容104的识别。作为一个示例,文档解析

器410可以响应于用户对文档104的选定部分的选择以及用户对识别文档104的选定部分的请求来评估文档102的内容104。

[0071] 作为该第三方面的第二变型,文档102可以由数学解析器204解析以标识数学表达式108(例如,文档102的展现出包括数学表达式108的高分类置信度206的区域302),并且文档的其余部分可以被分类为文本区域304。备选地,文档102可以由文本解析器202解析以标识文本表达式106(例如,文档102的展现出包括文本表达式106的高分类置信度206的区域302),并且文档的其余部分可以被分类为数学区域306。作为另一示例,文档102可以由文本解析器202和数学解析器204解析,以将区域302分别分类为文本区域304和数学区域306。在一些实施例中,文本解析器202和数学解析器204可以被同时应用于文档102,其中文档解析器410同时从文本解析器202和数学解析器204接收对各个区域302被分类为文本区域304和数学区域306中的一个区域的要求。在其他实施例中,解析器可以被顺序地应用;例如,可以首先将第一解析器应用于文档104(例如,更可能与内容104匹配的解析器),另一个解析器可以被应用于内容104的其余部分,该其余部分未被第一解析器肯定地分类。作为另一示例,分类可以是二进制的(例如,即使所得分类的分类置信度206为低,也基于分类置信度206将每个区域302分类为文本区域304和数学区域306中的一个区域),或者分类可能受到分类置信度阈值限制。即,各个解析器可以要求各种区域302提供与解析器的内容类型相匹配的内容104。歧义区域可以被保留为未分类(在提供附加内容104之后暂时或待重新评估),或者可以通过利用不同的解析技术(例如,将其与附近区域302分组在一起)而被解决,或者通过询问用户将内容104分类为文本区域304还是数学区域306而被解决。

[0072] 作为该第三方面的第三变型,解析可以以各种方式在文档102的内容104上迭代。作为第一个这样的示例,解析可以基于空间聚类将内容104分组为区域302,并且各个区域302可以相应地被表征。作为第二个这样的示例,解析可以涉及枚举文档102的符号,然后执行区域302的分类,直到符号已经被分类为文本区域304或数学区域306。例如,解析可以涉及首先选择文档102的符号的集合,然后将符号的集合映射到经识别的格式,诸如数学表达式108的数学表达式格式(例如,将特定的符号集合标识为与分数的格式匹配)或文本表达式106的句子和段落结构。

[0073] 作为该第三方面的第四变型,可以利用多种技术来执行解析和识别过程。作为第一个这样的示例,开发人员可以根据一组启发式(诸如特征性地表示文本表达式106和数学表达式108的一组符号和格式)来设计文本解析器202、数学解析器204、文本识别器114和/或数学识别器118。备选地或附加地,诸如人工神经网络或遗传算法之类的自适应技术可以被用来执行解析、分类和/或识别。例如,设备可以包括训练器,该训练器向包括文档解析器和文档识别器中的至少一个的自适应文档评估器提供训练文档,并根据训练文档来训练自适应文档评估器。然后(可选地根据由自适应技术生成的分类置信度阈值)可以由自适应文档评估器评估表达式的解析、区域302的分类和/或内容104的识别。例如,可以在作为文本表达式106和数学表达式108的一般示例而被提供的一组训练文档上和/或在诸如设备用户之类的特定个体的文档102上执行这样的训练,以使解析、分类和/或识别技术适应个体创建和/或编辑内容104的个人方式。许多这样的技术可以被用来根据本文提出的技术将解析、分类和识别技术应用于文档102。

[0074] E4. 解析和识别使用

[0075] 在本文提出的技术的实施例之中可以变化的第四方面涉及解析和/或识别任务的结果的使用。

[0076] 作为该第四方面的第一变型,设备可以以各种方式向用户指示解析、分类和/或识别的结果。作为第一个这样的示例,设备可以与数学区域306的内容104不同地可视地表示文本区域304的内容104,诸如至少部分地基于内容类型对区域302进行不同的颜色编码,和/或将不同的视觉样式应用于区域302,诸如以不同的视觉样式呈现指示文本区域304和数学区域306的边框。作为第二个这样的示例,设备可以以明确的方式向用户指示解析、分类和/或识别任务的完成和/或进度,诸如描述了文档102的评估中所涉及的各种处理任务的调用、进度、完成和/或结果的进度条、状态指示器和/或文本状态区域。还可以允许用户改变文档评估过程;例如,各个区域302的边界可以包括调整大小的控件,用户可以操纵该调整大小的控件来调节包括其中所包括的表达式在内的区域302的大小、形状和/或位置。还可以允许用户校正和/或重新指定区域302,诸如将区域302的分类从文本区域304改变为数学区域306(例如,因为分类不正确,或者因为用户意图对初始正确分类的区域302进行编辑,以提供与初始分类不一致的不同类型的内容104)。用户界面还可以允许用户创建新的区域302、移除现有区域302、合并或划分区域302和/或改变经识别的区域类型的集合(例如,用户可以针对特定类型的内容限定区域的新分类,诸如存储编程指令的代码区域,以便解析器和/或识别器可以自动标识存储代码表达式的区域302的内容104,将区域302分类为代码区域,并且应用适合于代码区域的用户界面)。

[0077] 作为该第四方面的第二变型,文档102的内容104的解析和识别可以影响文档102的文档的用户界面的呈现和适配,该用户界面编辑文档102的环境。当用户正在文档102的已被分类为数学区域306的区域302中创建和/或编辑内容104时,设备可以呈现用于数学工具的数学用户界面,诸如等式编辑器;并且当用户正在文档102的已被分类为文本区域304的区域302中创建和/或编辑内容104时,设备可以呈现用于文本工具的文本用户界面,诸如使用户能够选择粗体、下划线、字体样式和颜色以及拼写检查和语法检查功能的编辑控件。作为另一个示例,设备可以至少部分地基于解析、分类和/或识别来格式化和/或重新格式化文档102的内容104;例如,用户界面可以将文本区域304内的用户输入解释为文本并且可以应用以段落形式包装文本表达式106的流布局,并且可以将数学区域306内的用户输入解释为数学表达式108并且可以应用将等式按竖直顺序放置的行内表达式布局。

[0078] 作为该第四方面的第三变型,作为解析、分类和/或识别的结果,可以以各种方式存储文档102。作为第一个这样的示例,文档102可以被存储为文本表达式106和数学表达式108的聚合、拼贴或序列,这可以保留用户创建表达式的顺序以及按该顺序提供的任何语义意图(例如,包括数学表达式108和作为解释附加的文本表达式106的叙述)。作为第二个这样的示例,文档102可以被存储为文本表达式106和数学表达式108的分离的集合。分离的存储可以是有利的,例如以便在用户提供要应用于特定类型的内容104的搜索字符串的情况下(例如,请求搜索所有以字符序列“ \sin^2 ”为特征的数学表达式108)使得能够将分开的搜索过程应用于文本表达式106和数学表达式108。

[0079] 作为该第四方面的第四变型,可以基于各种类型的内容104提供各种补充工具。例如,当用户416选择文本区域304时,设备可以呈现包括以下的用户界面:一组文本表达式工具,诸如格式化工具;拼写和语法检查工具;以及翻译和叙述工具。当用户416选择数学区域

306时,设备可以利用包括以下的用户界面自动替换呈现文本表达式工具的用户界面:一组数学表达式工具,诸如数学表达编辑器;求解数学区域306中识别的数学表达式108的数学表达式求解器;以及将数学表达式108的结果可视化的可视化工具。

[0080] 图9是示例场景900的图示,其特征在于可以在创建、编辑和/或查看文档102期间向用户416呈现的用户界面。在该示例场景900中,向文档的各个区域302呈现区分文本区域304和数学区域306的视觉边界。附加地,在第一时间912处,用户416可以将光标902定位在区域302内,该区域302已被解析器分类为文本区域304,并且该设备可以呈现文本表达式用户界面904,该文本表达式用户界面904包括帮助用户416处理文本表达式106的文本表达式工具906的集合,该文本表达式工具906的集合诸如拼写检查工具906;语法检查工具906;向用户416读取文本表达式106的叙述工具906;以及将文本表达式106翻译成不同语言的翻译工具906。在第二时间914处,当用户416已经将光标902定位在区域302中时,该区域302已经被解析器分类为数学区域906,设备可以利用包括如下工具的数学表达式用户界面908自动替换文本表达式用户界面904,所述工具包括:表达式编辑器工具910,其帮助用户416输入数学表达式108(例如,插入数学表达式格式化结构,诸如指数和分数);数学表达式检查工具910,其验证由用户416输入的数学表达式108的准确性(例如,自动家庭作业评级工具);以及数学表达式可视化工具910,其自动生成并呈现数学区域306内的数学表达式108的可视化。当用户416稍后选择文本区域304时,设备可以利用文本表达式用户界面904替换数学表达式用户界面908。以这种方式,设备的用户界面可以响应于解析器的结果,自动适配以促进文档102中呈现的各种表达式的创建、编辑和查看,该解析器被应用以将区域302的内容102分类为文本区域304和数学区域306中的一个区域。可以设想出许多这样的技术,这些技术利用了根据本文提出的技术的解析结果。

[0081] F. 术语的使用

[0082] 尽管已经用特定于结构特征和/或方法动作的语言描述了主题,但是应该理解,所附权利要求书中限定的主题不必限于上述特定特征或动作。相反,上述特定特征和动作被公开为实现权利要求的示例形式。

[0083] 如在本申请中使用的,术语“组件”、“模块”、“系统”、“接口”等通常旨在指代与计算机有关的实体,硬件、硬件和/或硬件的组合、软件、或执行中的软件。一个或多个组件可以位于一台计算机上和/或可以被分布在两台或更多台计算机之间。

[0084] 此外,可以使用标准编程和/或工程技术以产生软件、固件、硬件或其任意组合来控制计算机以实现所公开的主题,从而将所要求保护的主体实现为方法、装置或制品。本文所使用的术语“制品”旨在涵盖可从任何计算机可读设备、载体或介质访问的计算机程序。当然,本领域技术人员将认识到可以在不背离所要求保护的主体范围或精神的情况下对该配置进行许多修改。

[0085] 本文提供了实施例的各种操作。在一个实施例中,所描述的操作中的一个或多个可以构成存储在一个或多个计算机可读介质上的计算机可读指令,该指令如果由计算设备执行,则将使计算设备执行所描述的操作。描述一些或所有操作的顺序不应解释为暗示这些操作必定与顺序有关。受益于此描述的本领域技术人员将意识到替代的排序。此外,将理解的是,并非在本文提供的每个实施例中必须呈现所有操作。

[0086] 本文中被描述为“示例”的任何方面或设计不必被解释为优于其他方面或设计。而

是,词语“示例”的使用旨在呈现可能与本文提出的技术有关的一个可能的方面和/或实现。这样的示例对于这样的技术不是必需的也不旨在是限制的。这样的技术的各种实施例可以单独地或与其他特征组合地包括这样的示例,和/或可以改变和/或省略所例示的示例。

[0087] 如在本申请中使用的,术语“或”旨在意指包括性的“或”而不是排他性的“或”。也就是说,除非另有说明或从上下文可以清楚看出,否则“X采用A或B”旨在意指任何自然的包含性枚举。也就是说,如果X采用A;X采用B;或X采用A和B两者,那么在任意前述实例中都满足“X采用A或B”。另外,除非另有说明或从上下文清楚地指向单数形式,否则在本申请和所附权利要求书中使用的冠词“一”和“一个”通常可以被解释为意指“一个或多个”。

[0088] 而且,尽管已经相对于一个或多个实现示出和描述了本公开,但是基于对本说明书和附图的阅读和理解,本领域的其他技术人员将想到等同的变型和修改。本公开包括所有这样的修改和变型,并且仅由所附权利要求的范围限制。特别是关于由上述组件(例如,元素、资源等)执行的各种功能,除非另有说明,否则用于描述此类组件的术语旨在与执行所述组件的指定功能的任何组件相对应。(例如,其在功能上是等效的)——即使在结构上不等同于在本公开的本文例示的示例实现中执行功能的所公开的结构。另外,尽管可能仅仅关于若干实现中的一个实现公开了本公开的特定特征,但是这种特征可以与其他实现的一个或多个其他特征组合,其对于任何给定的或特定的应用而言可能期望和有利的。此外,就在详细描述或权利要求中使用术语“包括”、“具有”、“带有”或其变型的程度而言,这些术语旨在以类似于术语“包括”的方式被包括在内。

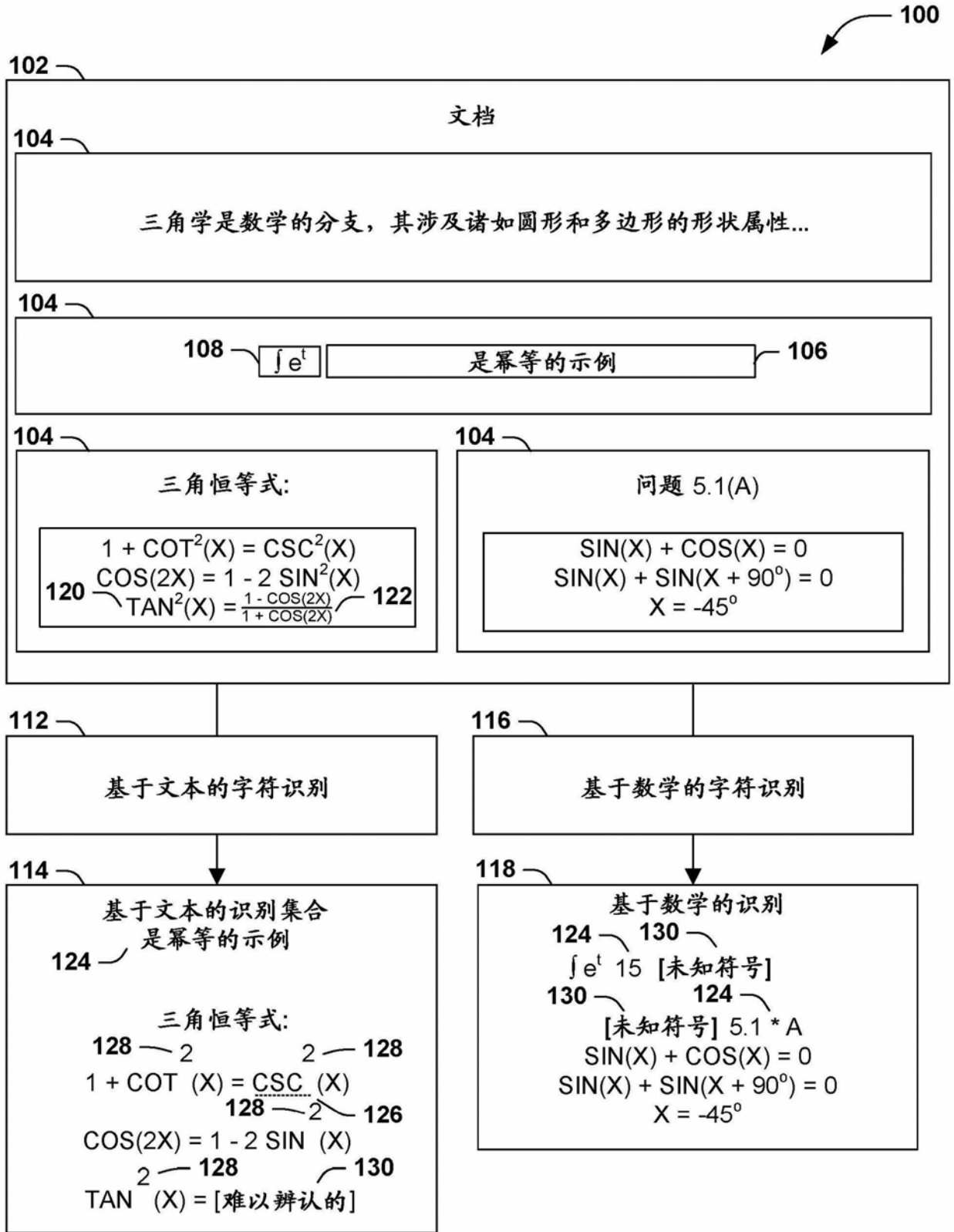


图1

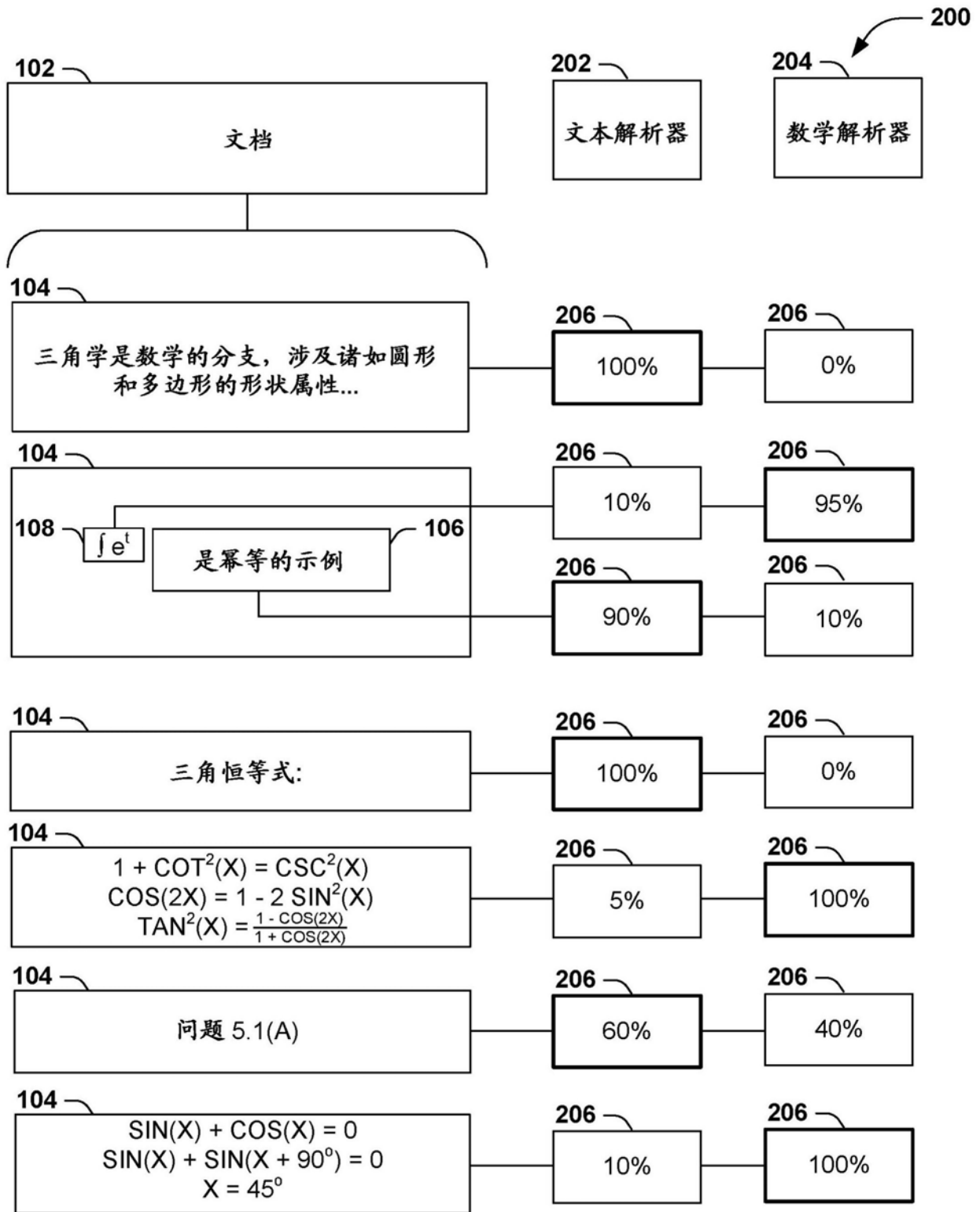


图2

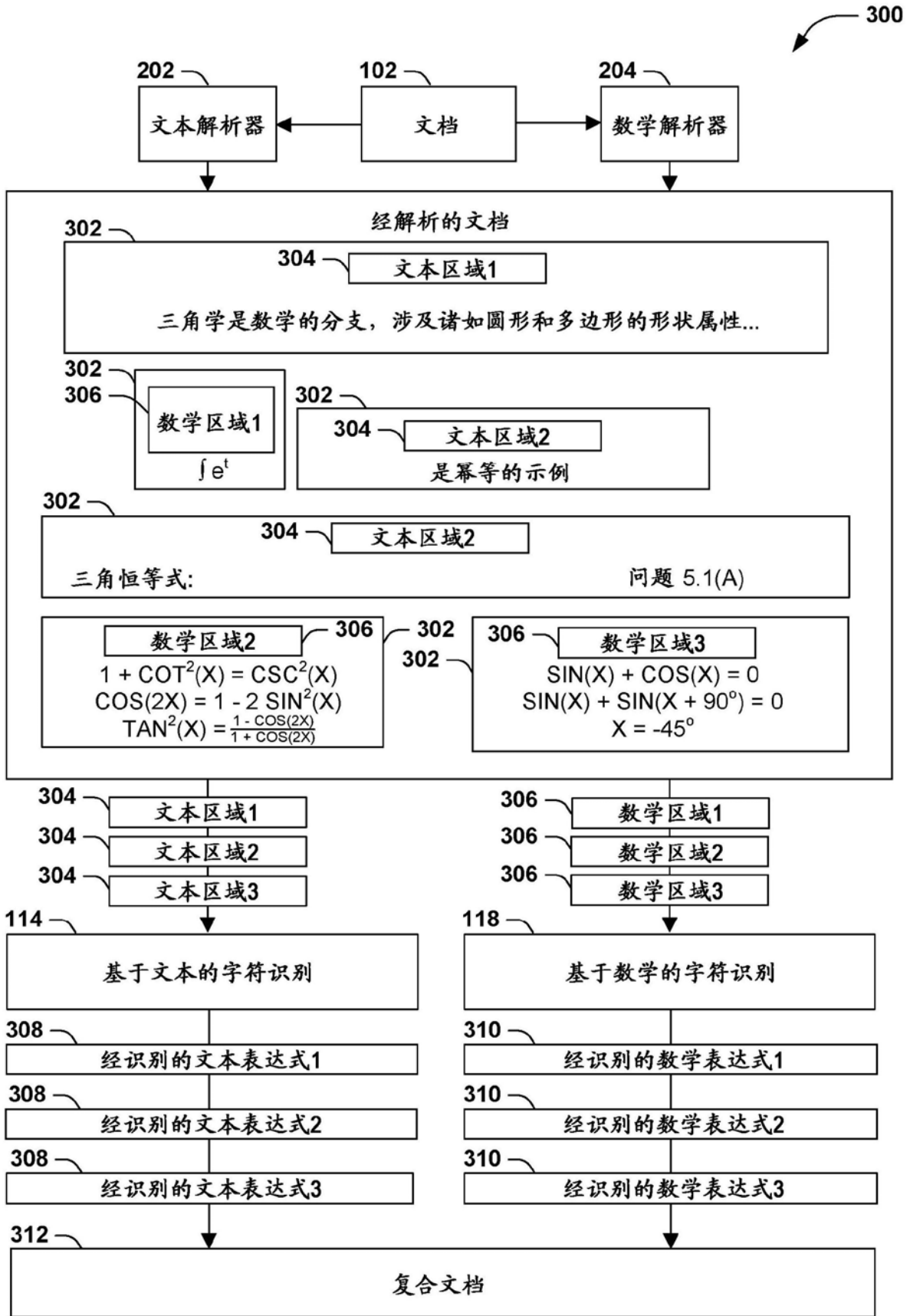


图3

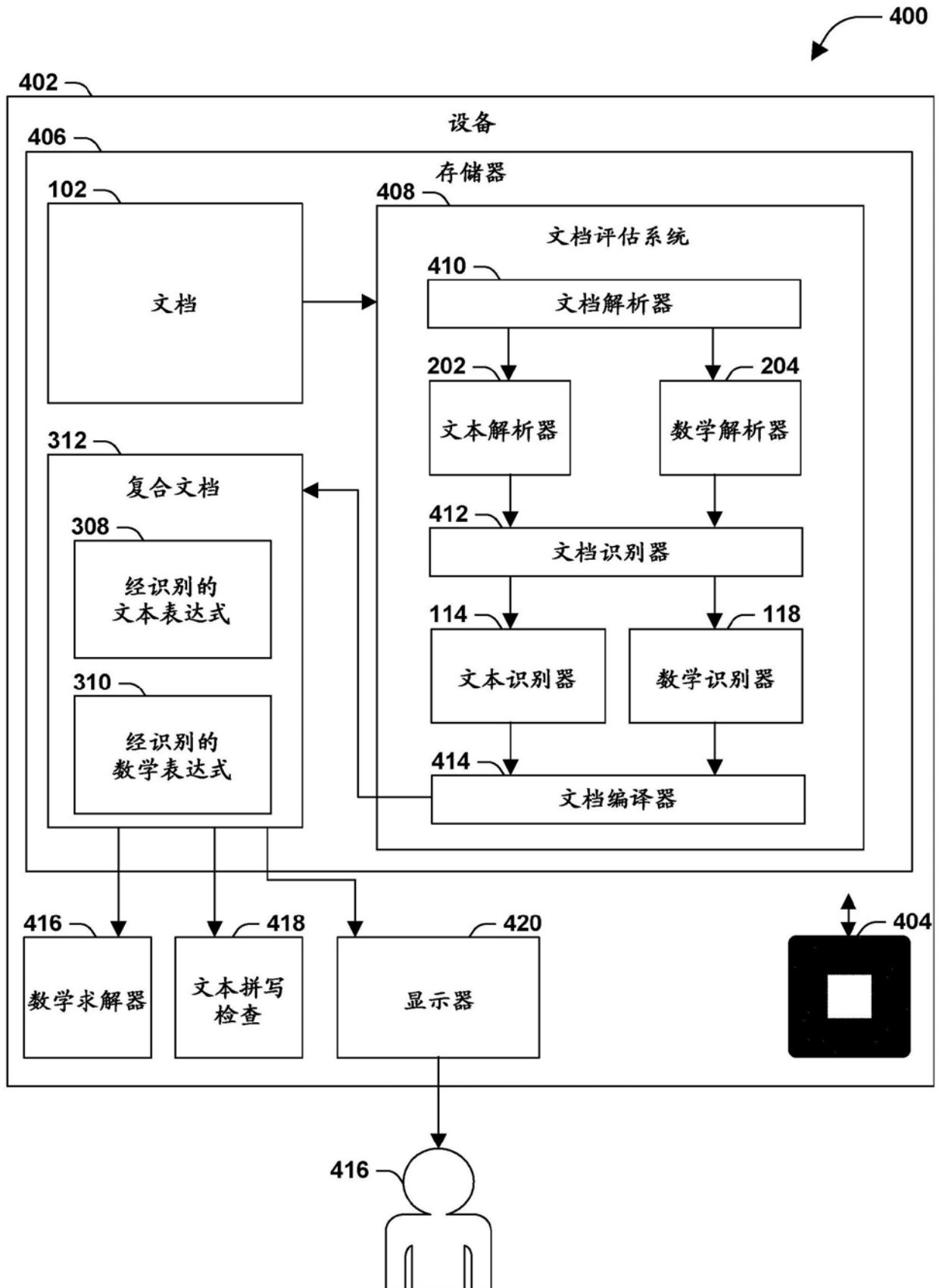


图4

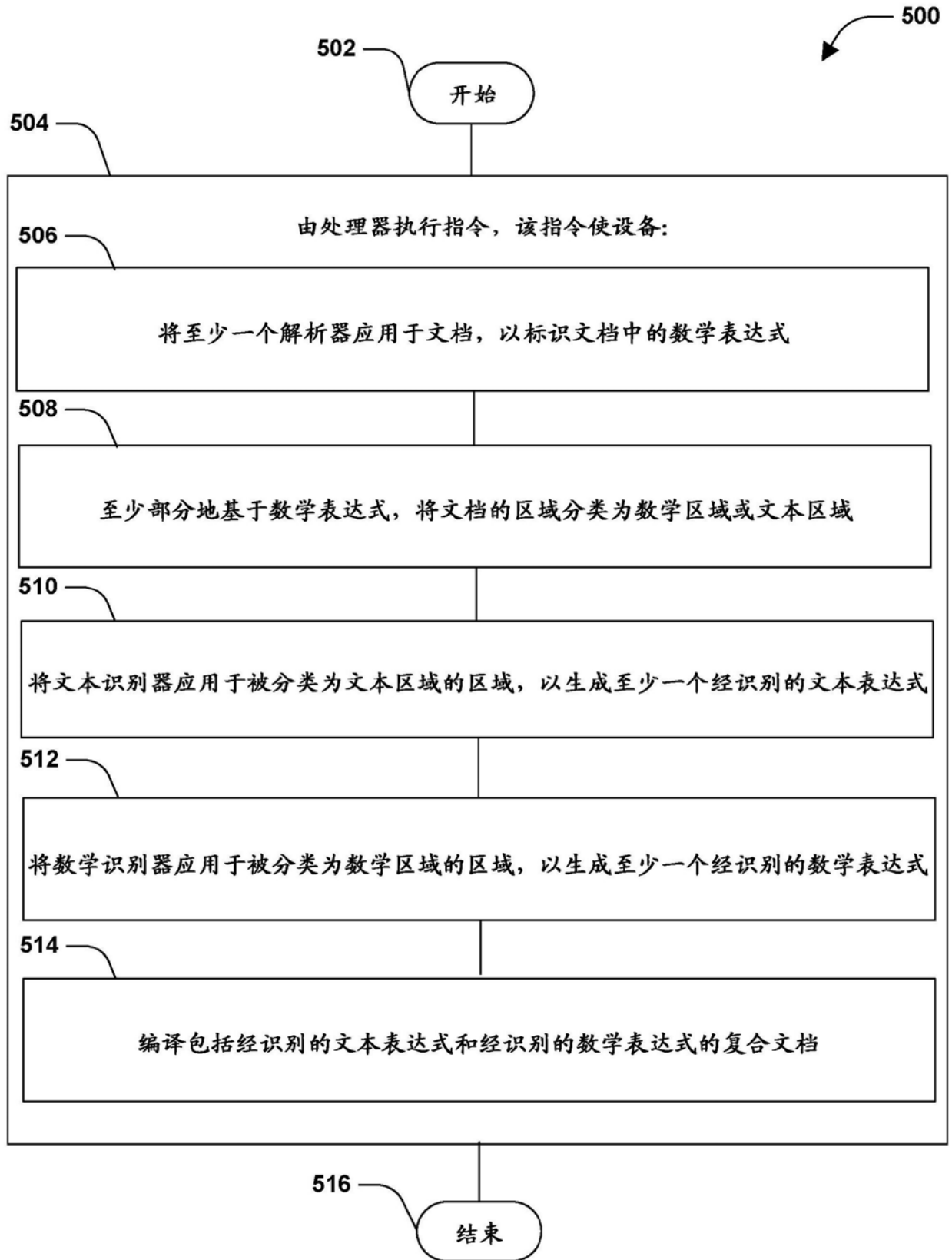


图5

600

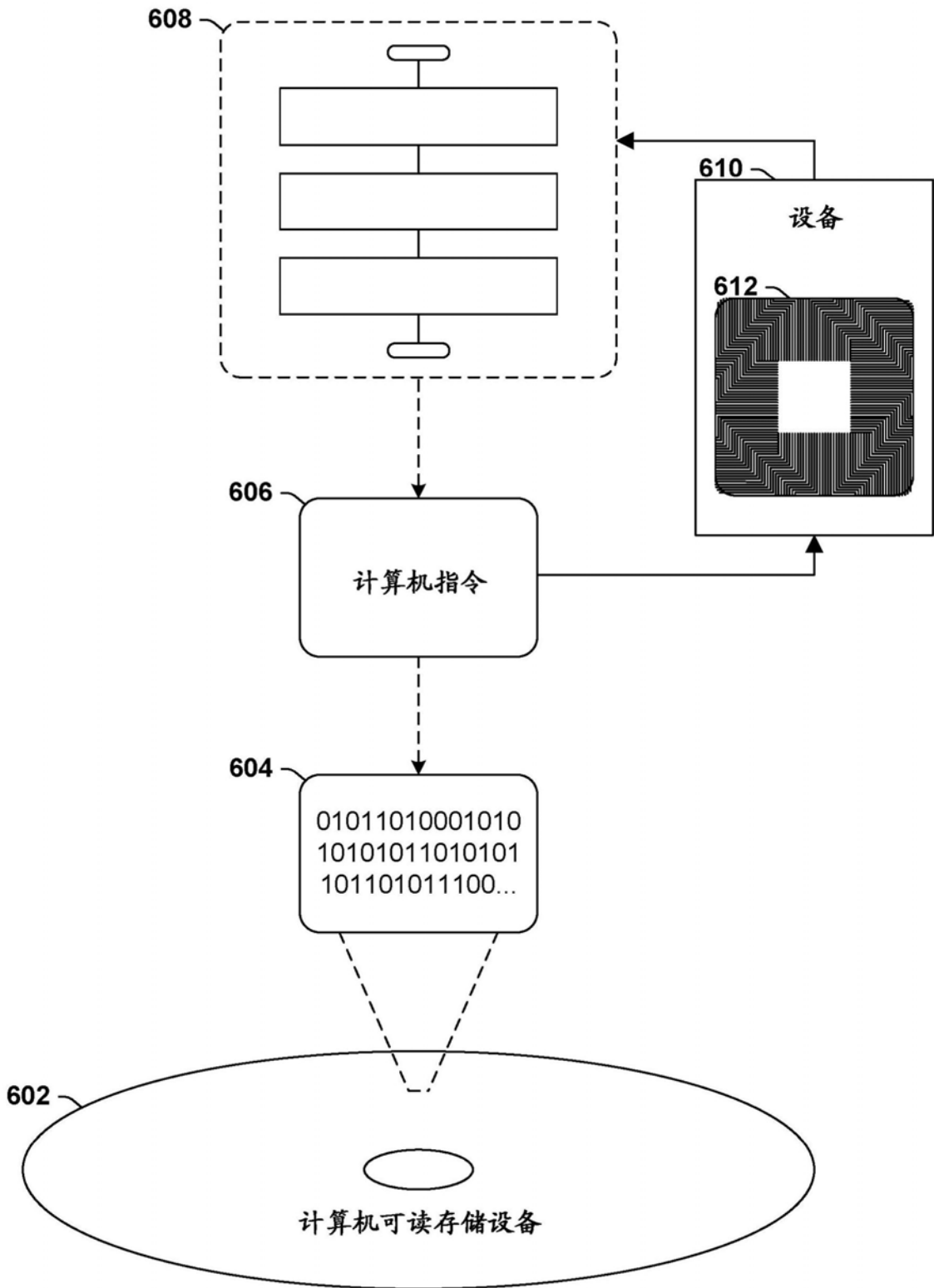


图6

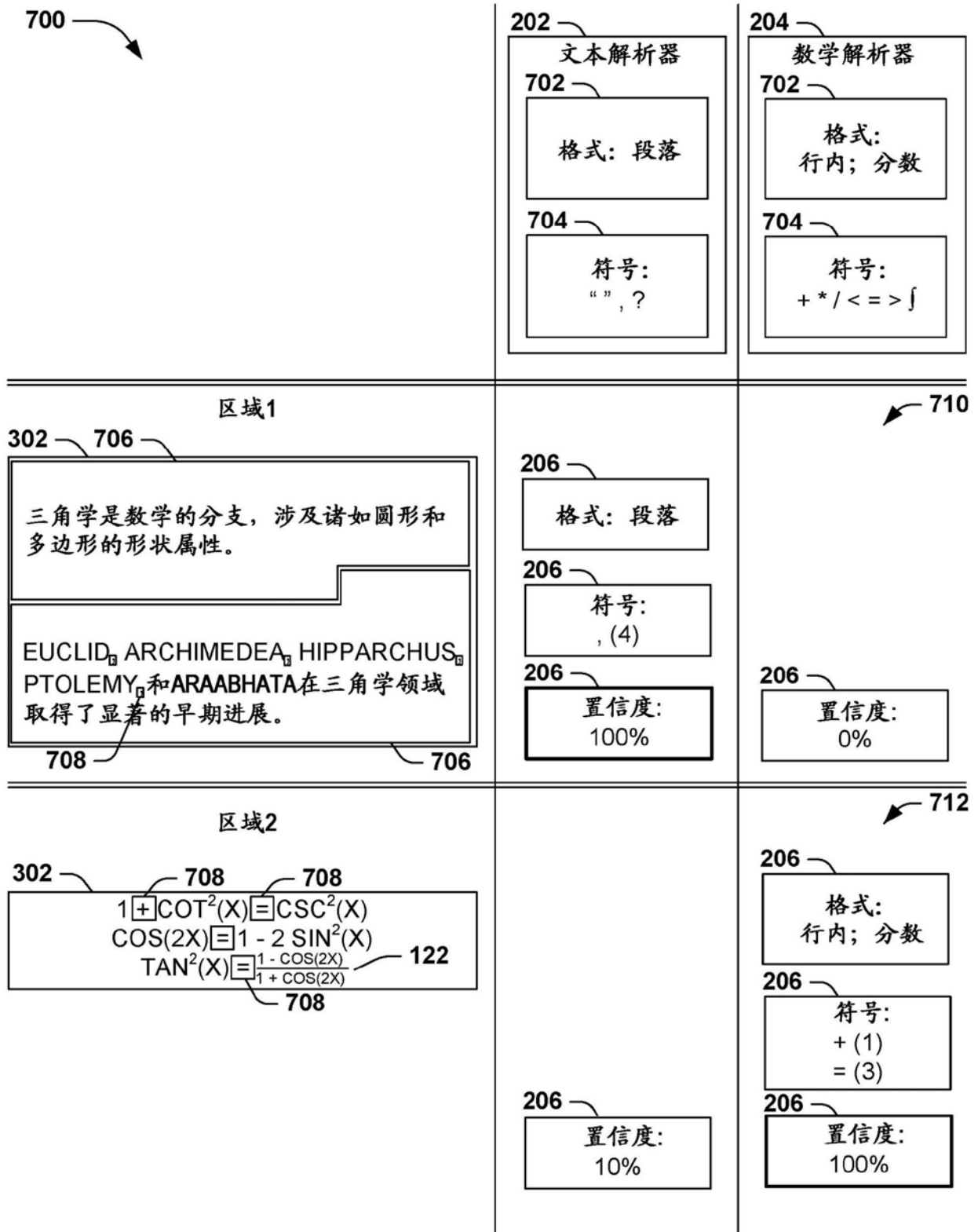


图7

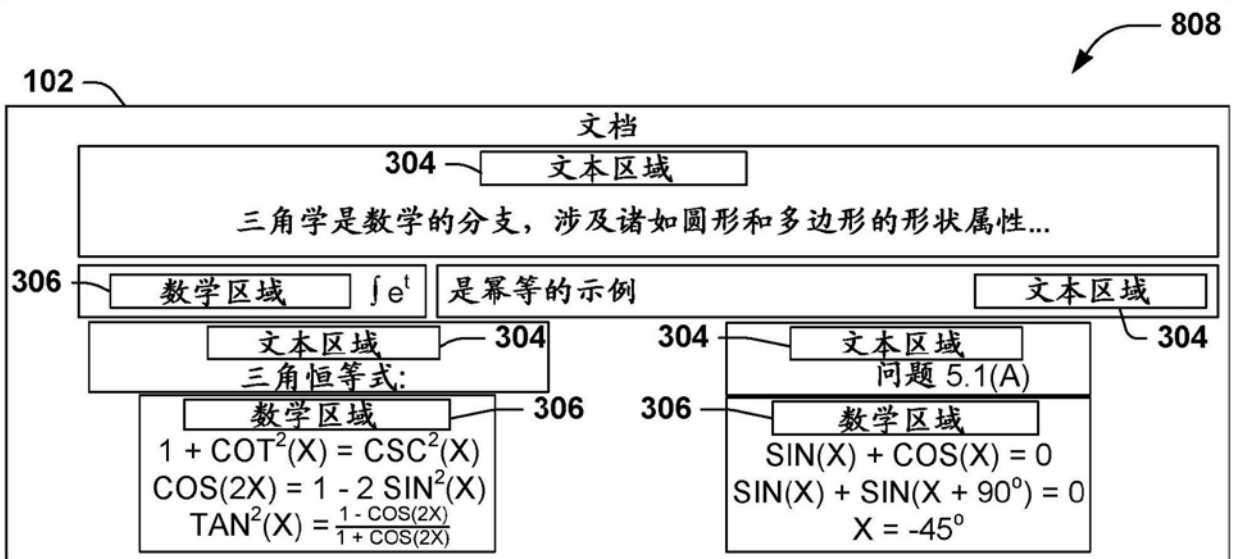
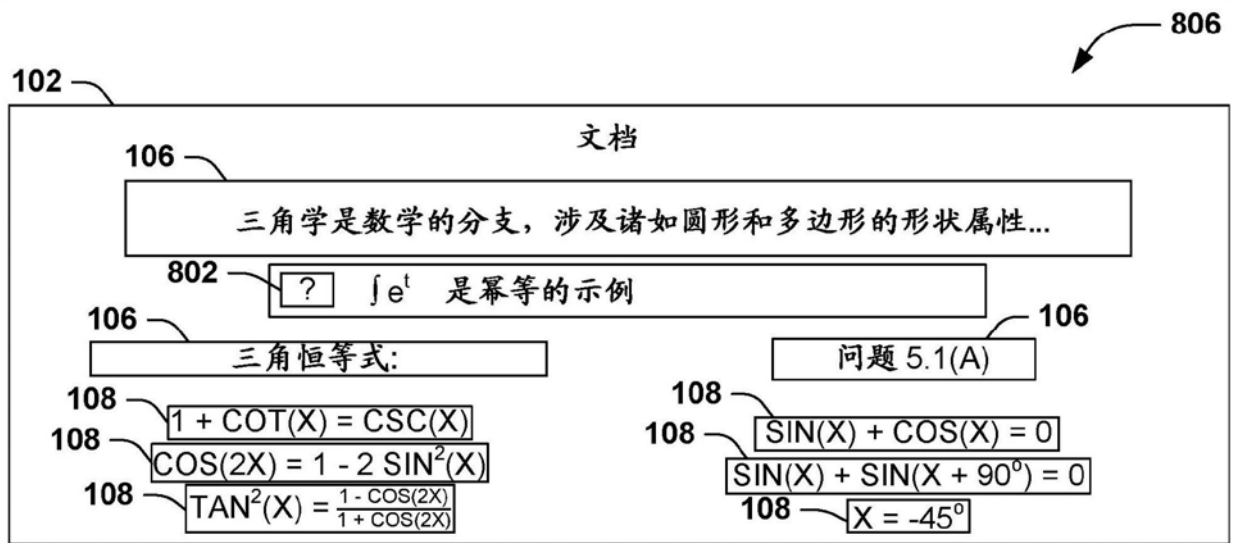
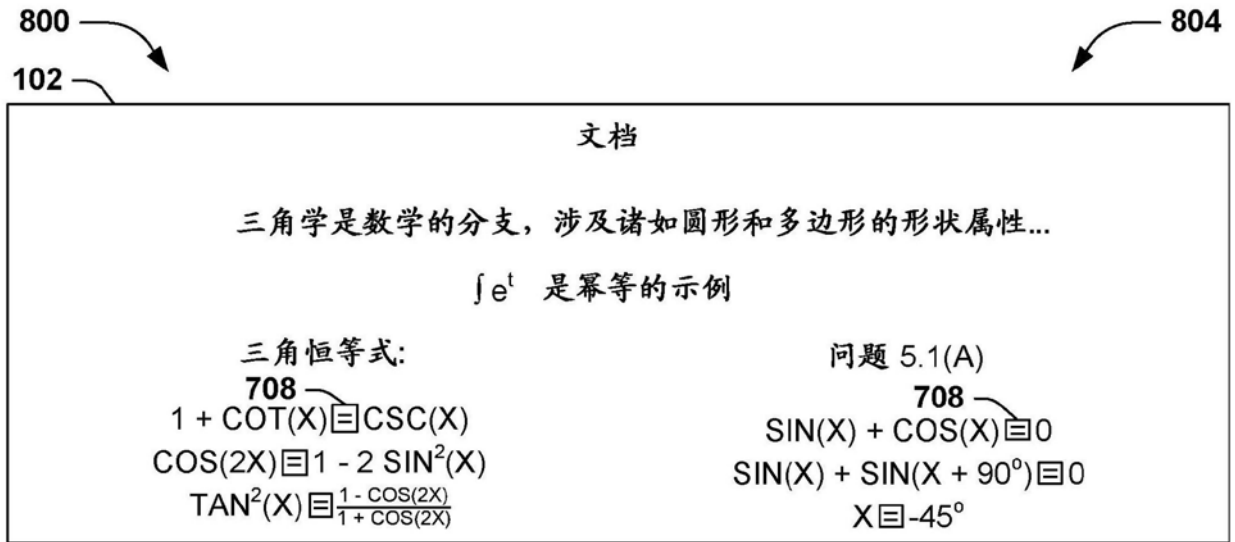


图8

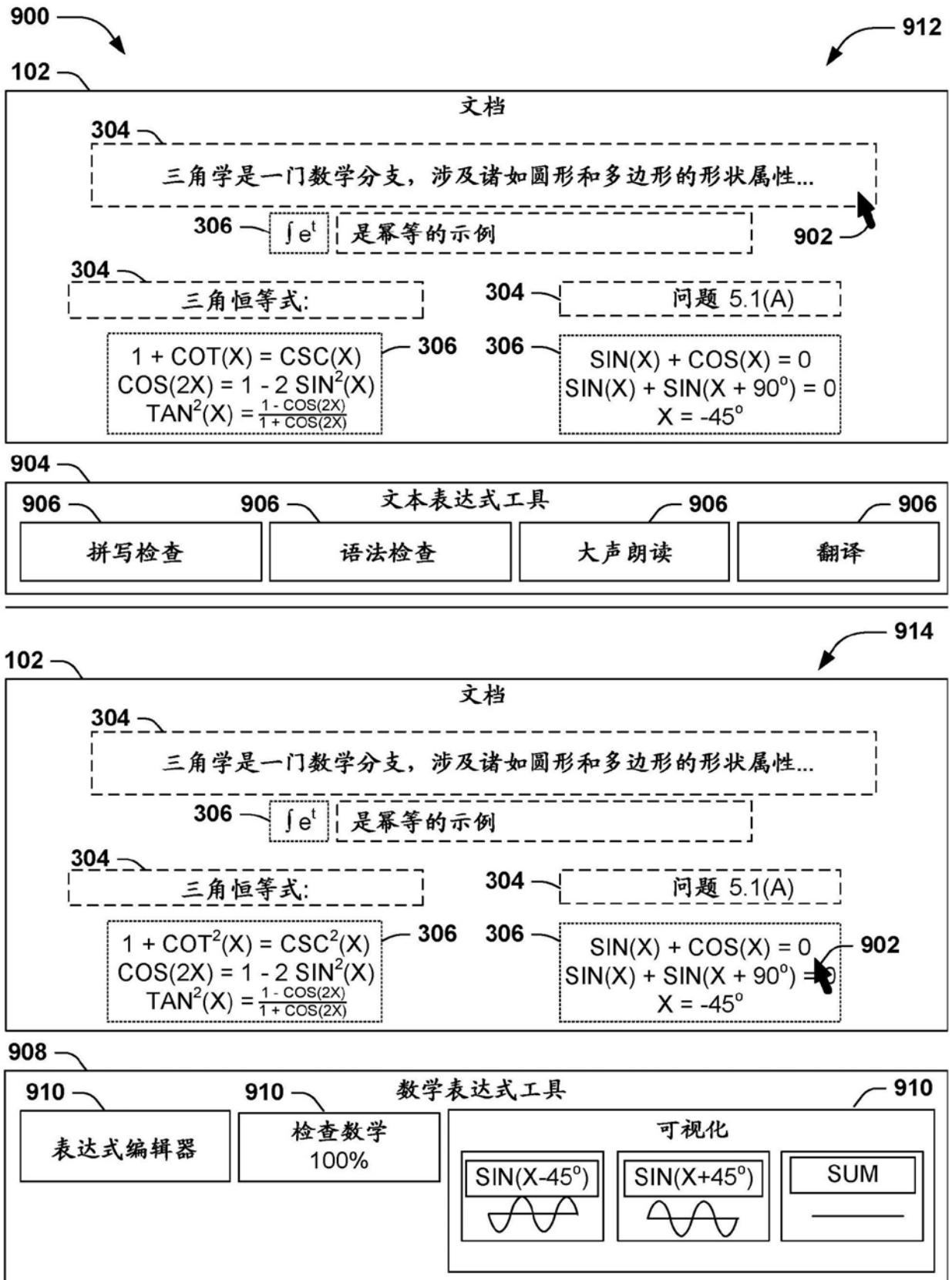


图9