(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101425132 B
(45)授权公告日 2012.10.17

(21)申请号 200810170687.8

(22)申请日 2008.10.30

(30)优先权数据
2007-284294 2007.10.31 JP

(73)专利权人 富士通株式会社
地址 日本神奈川县川崎市

(72)发明人 田中宏 藤井勇作 武部浩明
堀田悦伸 皆川明洋 小泽宪秋
藤本克仁 深沢克夫 稻见真树
渡边起一郎

(74)专利代理机构 北京三友知识产权代理有限公司 11127
代理人 李辉

(51)Int.Cl.
G06K 9/00(2006.01)

(56)对比文件
CN 1225484 A,1999.08.11, 全文.
US 7149970 B1,2006.12.12, 全文.
CN 1215201 A,1999.04.28, 全文.
JP 特开 2007-222747 A,2007.09.06, 说明书第 0011-0024 段,附图 1,2,4.
WO 2007/094684 A2,2007.08.23, 全文.

审查员 高霖

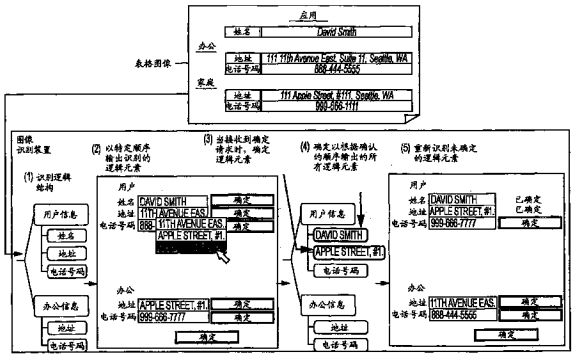
权利要求书 1 页 说明书 16 页 附图 23 页

(54)发明名称

图像识别装置和识别方法及记录图像识别程序的存储介质

(57)摘要

本发明提供图像识别装置和方法及记录图像识别程序的存储介质。图像识别装置对其中字符串被描述为逻辑元素的图像中的字符串与组成逻辑结构的逻辑元素之间的对应关系进行识别以识别每个逻辑元素。该图像识别装置包括:输出装置,用于在识别或重新识别所述对应关系时,输出所识别的逻辑元素;第一确定装置,用于在从用户接收到用于确定所述逻辑元素的确定请求的输入时,确定特定逻辑元素为正确;第二确定装置,用于确定在第一确定装置确定的所述逻辑元素之前输出的并根据用户的确认而定位的所有逻辑元素的正确性;以及重新识别装置,用于基于每个逻辑元素的确定内容,来重新识别仍未被确定为正确的逻辑元素与字符串之间的对应关系。



CN 101425132 B

1. 一种图像识别装置，所述图像识别装置对图像中的字符串与组成逻辑结构的多个逻辑元素之间的对应关系进行识别以识别每个逻辑元素，在所述图像中，所述字符串被描述为所述多个逻辑元素，所述图像识别装置包括：

输出装置，其用于在识别或重新识别所述对应关系时，输出所识别出的被描述为所述多个逻辑元素的字符串；

第一确定装置，其用于在从识别出所述输出装置输出的识别结果的用户接收到用于确定逻辑元素的确定请求的输入时，确定所述多个逻辑元素中的特定逻辑元素的正确性；

第二确定装置，其用于确定在通过所述第一确定装置确定的逻辑元素之前输出的并根据所述用户进行的确认而定位的所有逻辑元素的正确性；以及

重新识别装置，其用于基于所述第一确定装置和所述第二确定装置确定为正确的每个逻辑元素的确定内容，来重新识别所述多个逻辑元素中仍未被确定为正确的逻辑元素和所述字符串之间的对应关系。

2. 根据权利要求 1 所述的图像识别装置，所述图像识别装置还包括：

存储装置，其用于存储由所述重新识别装置识别的识别结果，

其中，当所述重新识别装置输出所述重新识别结果时，所述输出装置针对每个逻辑元素，将所述存储装置存储的识别结果与当前显示的识别结果进行比较，并针对其中所述存储装置存储的识别结果与当前显示的识别结果不一致的逻辑元素，输出用于指示所述重新识别装置的重新识别结果与当前显示的识别结果不一致的信息。

3. 一种图像识别方法，所述图像识别方法对图像中的字符串与组成逻辑结构的多个逻辑元素之间的对应关系进行识别以识别每个逻辑元素，在所述图像中，所述字符串被描述为所述多个逻辑元素，所述图像识别方法包括以下步骤：

输出步骤，其在识别或重新识别所述对应关系时，输出所识别出的被描述为所述多个逻辑元素的字符串；

第一确定步骤，其在从识别出所述输出步骤输出的识别结果的用户接收到用于确定逻辑元素的确定请求的输入时，确定所述多个逻辑元素中的特定逻辑元素为正确的；

第二确定步骤，其确定在所述第一确定步骤确定的逻辑元素之前输出的并根据所述用户进行的确认而定位的所有逻辑元素的正确性；以及

重新识别步骤，其基于所述第一确定步骤和所述第二确定步骤确定为正确的每个逻辑元素的确定内容，来重新识别所述多个逻辑元素中仍未被确定为正确的逻辑元素与所述字符串之间的对应关系。

4. 根据权利要求 3 所述的图像识别方法，所述图像识别方法还包括以下步骤：

存储步骤，其存储通过所述重新识别步骤识别的识别结果，

其中，当输出所述重新识别步骤的重新识别结果时，所述输出步骤针对每个逻辑元素，将所述存储步骤存储的识别结果与当前显示的识别结果进行比较，并针对其中所述存储步骤存储的识别结果与当前显示的识别结果不一致的逻辑元素，输出用于指示所述重新识别步骤的重新识别结果与当前显示的识别结果不一致的信息。

图像识别装置和识别方法及记录图像识别程序的存储介质

技术领域

[0001] 本发明涉及图像识别装置、图像识别方法，及记录图像识别程序的存储介质。

发明内容

[0002] 根据本发明的图像识别装置对图像中的字符串与组成逻辑结构的逻辑元素之间的对应关系进行识别，其中所述字符串被描述为多个所述逻辑元素以识别每个逻辑元素。所述图像识别装置包括：输出装置，其用于在识别或重新识别所述对应关系时，输出所识别的多个逻辑元素；第一确定装置，其用于在从识别出所述输出装置输出的所述识别结果的用户接收到用于确定接收到所述逻辑元素的确定请求的输入时，所述第一确定装置用于确定所述多个逻辑元素中的特定逻辑元素是正确的；第二确定装置，其用于确定在所述第一确定装置确定的所述逻辑元素之前输出的并根据所述用户进行的确认而定位的所有所述逻辑元素的正确性；以及重新识别装置，其用于基于所述第一确定装置和所述第二确定装置确定为正确的每个逻辑元素的确定内容，来重新识别仍未被确定为正确的逻辑元素与所述字符串之间的对应关系。

附图说明

[0003] 图 1 是用于描述根据第一实施方式的图像识别装置的概况和特征的图；

[0004] 图 2 是示出根据第一实施方式的图像识别装置的结构的实施例的框图；

[0005] 图 3 图示表格图；

[0006] 图 4 是用于描述识别结果显示屏面的图；

[0007] 图 5A 和 5B 是用于描述逻辑结构模型存储部件的图；

[0008] 图 6 是用于描述识别结果显示缓冲器的图；

[0009] 图 7 是用于描述纠错操作的图；

[0010] 图 8A 到 8D 是用于描述逻辑元素的自动确定的图；

[0011] 图 9A 到 9B 是用于描述（转换后的）逻辑结构存储部件的图；

[0012] 图 10A 和 10B 是用于描述纠错操作的图；

[0013] 图 11 是示出根据第一实施方式的图像识别装置执行的处理的实施例的流程图；

[0014] 图 12 是用于描述另一识别结果显示屏面的图；

[0015] 图 13A 和 13B 是用于描述根据第二实施方式的图像识别装置的概况和特征的图；

[0016] 图 14 是示出根据第二实施方式的图像识别装置的结构的实施例的框图；

[0017] 图 15A 和 15B 是用于描述识别结果显示缓冲器和重新识别结果缓冲器的图；

[0018] 图 16 是示出根据第二实施方式的图像识别装置执行的处理的实施例的流程图；

[0019] 图 17A 和 17B 是用于描述另一识别结果显示屏面的图；

[0020] 图 18A 和 18B 是用于描述根据第三实施方式的图像识别装置的概况和特征的图；

[0021] 图 19A 和 19B 是用于描述识别结果显示缓冲器和重新识别结果缓冲器的图；

[0022] 图 20 是示出根据第三实施方式的图像识别装置执行的处理的实施例的流程图；

[0023] 图 21A 和 21B 是用于描述另一识别结果显示屏面的图；以及

[0024] 图 22 是示出执行图像识别程序的计算机的结构的实施例的框图。

具体实施方式

[0025] 近年来，商业的计算机化伴随着对各种类型的计算机化文档的使用。因此，增加了将纸质文档转换成计算机化文档的图像识别技术的重要性。图像识别技术包括光学字符读取器 (OCR) 或光学字符识别 (OCR)。

[0026] 然而，通过图像识别技术进行的文档识别可能在识别结果中产生错误。在这种情况中，用户通过从视觉上对输入的文档中的字符串与识别结果中的字符串进行比较，来纠正识别结果中的错误。

[0027] 日本特开平 No. 5-158605 公开了一种基于在数据输入装置中输入的数据项来限制要输入的数据项的技术，在所述数据输入装置中，用户按顺序输入这些数据项。

[0028] 上面描述的相关领域中的技术具有纠正操作的负荷变得更重的问题。例如，假设在处理因客户而不同或者因交易而不同的大量表格的银行中执行上面的纠正操作，则这种纠正操作可能耗费许多工时并导致增加的工作负荷。

[0029] 实现下面描述的本发明的实施方式以解决上面的问题并减小纠正操作的负荷。

[0030] 根据本发明实施方式的图像识别装置防止未确定的逻辑元素（即仍未确定为正确的逻辑元素）变得更差。

[0031] 这里将参照附图详细描述本发明的实施方式。现在将描述这些实施方式中使用的主要术语、根据第一实施方式的图像识别装置的概况、根据第一实施方式的图像识别装置的结构，以及根据第一实施方式的图像识别装置中的处理，随后再描述其他实施方式。

[0032] 第一实施方式

[0033] 术语的描述

[0034] 现在将描述本发明的实施方式中使用的主要术语。这些实施方式中使用的“图像”意指其中字符串被描述为组成特定逻辑结构的多个逻辑元素的图像。例如，“图像”可以意指表格图像。假设表格图像包括其中输入“姓名”、“地址”以及“电话号码”作为用户信息的项，和其中输入“地址”和“电话号码”作为办公信息的项，并且假设字符串被表述在表格图像中的每个项中。在这种表格图像中，“逻辑结构”意指这样的结构，在所述结构中，“姓名”、“地址”以及“电话号码”位于“用户信息”下方，“地址”和“电话号码”位于“办公信息”下方，而“用户信息”位于和“办公信息”相同的层次中。“姓名”、“地址”以及“电话号码”等对应于组成“逻辑结构”的“逻辑元素”。换言之，例如，“逻辑结构”被设置为由多个项有意或无意形成的结构，并且当表格图像的格式的创建者创建该表格图像的格式时，所述多个项被设置为“逻辑元素”。当图像识别装置处理该表格图像时，该表格图像中由多个项形成的结构被处理为“逻辑结构”，而所述多个项被处理为“逻辑元素”。“图像”并不限于表格图像，而可以是任何图像，例如问卷图像，只要在字符串被表述在每个项中即可。

[0035] 根据第一实施方式的“图像识别装置”识别字符串与逻辑元素之间的对应关系，以识别这种图像中的每个逻辑元素。例如，“图像识别装置”通过诸如 OCR 的图像识别技术识别表格图像中描述的每个字符串。此外，“图像识别装置”通过使用例如逻辑结构模型来识别每个字符串被描述为哪种逻辑元素。例如，“图像识别装置”通过使用诸如 OCR 的图像识

别技术来识别被手写为“David Smith”的字符串“David Smith”，并识别出字符串“David Smith”被描述为逻辑元素“姓名”，以识别表格图像中的每个逻辑元素。在下面描述的这些实施方式中，这种识别被适当地称为“逻辑结构识别”或“识别逻辑结构”。

[0036] 然而，“图像识别装置”并不总是正确地识别字符串与逻辑元素之间的对应关系，而可能获得错误的识别结果。“图像识别装置”可能错误地识别表格图像的各个逻辑元素。例如，“图像识别装置”可能将手写地址“11111th Avenue East，Suite11，Seattle，WA”错误地识别为“77711thAvenue East，Suite11，Seattle，WA；”，或者“图像识别装置”可能将针对“office（办公）”地址写的地址“11111th Avenue East，Suite11，Seattle，WA”错误地识别为“home（家庭）”地址。在这种情况中，执行纠正操作来纠正识别结果中的任何错误。提供根据第一实施方式的“图像识别装置”用于减小纠正操作的负荷。

[0037] 根据第一实施方式的图像识别装置的概况

[0038] 现在将参照图 1 描述根据第一实施方式的图像识别装置的概况。图 1 是描述根据第一实施方式的图像识别装置的概况的图。

[0039] 如上文所述，根据第一实施方式的图像识别装置识别图像中字符串与多个逻辑元素之间的对应关系，以识别每个逻辑元素从而减小纠正操作的负荷，在所述图像中，字符串被描述为组成特定逻辑结构的逻辑元素。

[0040] 例如，根据第一实施方式的图像识别装置识别图 1 中上半部分中示出的表格图像内的每个逻辑元素。具体的是，该图像识别装置识别该表格图像中字符串与多个逻辑元素之间的对应关系，在所述表格图像中，作为用户信息的字符串“姓名”、作为办公信息的字符串“地址”和“电话号码”，以及作为用户信息的字符串“地址”和“电话号码”被描述为逻辑元素。

[0041] 如图 1 中下半部分中所示，根据第一实施方式的图像识别装置存储被表示为图 1 中树结构的概念的逻辑结构模型，并且基于该逻辑结构模型来识别逻辑结构。具体的是，图像识别装置存储这样的逻辑结构模型，所述逻辑结构模型被设置成对应于其中“姓名”、“地址”以及“电话号码”位于“用户信息”下方，“地址”和“电话号码”位于办公信息的下方，而“用户信息”位于和“办公信息”相同的层次内的结构。在该逻辑结构模型中还可设置位置关系，例如其中在“用户信息”下方描述“办公信息”。

[0042] 在上面的结构中，当根据第一实施方式的图像识别装置识别表格图像（参照图 1 中的 (1)）的逻辑结构时，该图像识别装置以特定顺序（参照图 1 中的 (2)）将作为识别结果的多个逻辑元素输出到输出单元。例如，图像识别装置按照下面的顺序将作为用户信息中的“姓名”的“DavidSmith”、作为用户信息中的“地址”的“11th Avenue East，Suite11，Seattle，WA”、作为用户信息中的“电话号码”的“888-444-5555”，作为办公信息中的“地址“的”Apple Street，#111，Seattle，WA”，以及作为办公信息中的“电话号码”的“999-666-7777”输出到输出单元。如从图 1 中上半部分中示出的表格图像（应用）的内容可见，图像识别装置错误地颠倒了用户信息中的“地址”和“电话号码”与办公信息中的“地址”和“电话号码”。

[0043] 当图像识别装置从确认识别结果的用户处接收到用于确定多个逻辑元素中的特定逻辑元素正确的确定请求的输入时，图像识别装置确定该特定逻辑元素是正确的（参照图 1 中的 (3)）。例如，当图像识别装置从用户处接收到确定请求的输入，从而确定用户信息

中的“地址”不是“11thAvenue East,Suite11,Seattle,WA " 而是 " Apple Street,#111,Seattle,WA”时,图像识别装置确定该“地址”正确。

[0044] 然后,图像识别装置确定在确定的特定逻辑元素之前被输出、并且根据用户的确认(参照图 1 中的 (4)) 而定位的所有逻辑元素的正确性。例如,假设用户通常在图 1 中的 (2) 示出的输出中从顶至底顺序地逐一确认这些项。在这种情况中,在确定的逻辑元素(用户信息中的“地址”)之前被输出、并且根据用户的确认而定位的逻辑元素意指定位在用户信息中的“地址”上方的逻辑元素。用户信息中的“姓名”对应于在确定的逻辑元素之前输出的逻辑元素,并且对应于根据用户的确认被定位的逻辑元素。因此,图像识别装置还确定用户信息中的“姓名”是正确的。

[0045] 根据第一实施方式,用确定的字符串替换图像识别装置中存储的逻辑结构模型中被确定要纠正的逻辑元素,如图 1 中的 (3) 和 (4) 所示,并且转换逻辑结构模型本身。

[0046] 然后,根据第一实施方式的图像识别装置基于针对要纠正的逻辑元素的确定内容,来重新识别仍未被确定为正确的逻辑元素的逻辑结构。因此,通过在图像识别装置中以特定顺序输出重新识别结果来给出图 1 中的 (5)。例如,图像识别装置基于转换的逻辑结构模型来识别字符串与逻辑元素之间的对应关系。具体来说,图像识别装置重新识别用户信息中的“电话号码”和办公信息中的“地址”和“电话号码”,以获得“999-666-7777”、“11th Avenue East,Suite11,Seattle,WA,”以及“888-444-5555”作为重新识别的结果。如从图 1 中 (2) 的输出和图 1 中 (5) 的输出之间的比较所见,(2) 中的输出与 (5) 中的输出在用户信息内的“电话号码”和办公信息内的“地址”和“电话号码”方面不同。换言之,重新识别的结果中反映出确定的内容。

[0047] 根据第一实施方式的图像识别装置以上面方式可以减小纠正操作的负荷。在图 1 中示出的实施例中通过图像识别装置进行的第一识别的结果中,错误地识别了用户信息中除了“姓名”以外的所有信息,并且需要用户纠正或者确定所有被错误识别的项。然后,当确定用户信息中的“地址”为正确时,根据第一实施方式的图像识别装置确定不仅用户信息中的“地址”为正确,而且其中的“姓名”为正确,并且重新识别这些信息。因此,能够减少用户必须按顺序确定的项的数目。此外,例如,用户信息中的“地址”和其中的“电话号码”之间的不一致表明用户信息中的“电话号码”不正确,并且用户信息中的“地址”和办公信息中的“地址”的不一致表明办公信息不正确。因此,在某些情况中,可以减少或者甚至消除用户必须按顺序纠正的项的数目。

[0048] 如上面所描述的,当用户确定或纠正特定逻辑元素时,根据第一实施方式的图像识别装置参考具有固定的确定逻辑元素的逻辑结构模型,并立即重新识别逻辑结构以改善其他未确定逻辑元素的识别结果。如果针对未确定逻辑元素的识别结果的候选数被减小为 1,则能够确定该逻辑元素。因此,因为用户确定或纠正的结果可以被用来确定其他逻辑元素,所以能够减小用户的纠正操作的负荷。

[0049] 根据第一实施方式的图像识别装置的结构

[0050] 现在将参照图 2 到 10 描述根据第一实施方式的图像识别装置。图 2 是示出根据第一实施方式的图像识别装置的结构的实施例的框图。图 3 例示了表格图像。图 4 是用于描述识别结果显示屏面的图。图 5A 和 5B 是用于描述逻辑结构模型存储部件的图。图 6 是用于描述识别结果显示缓冲器的图。图 7 是用于描述纠错操作的图。图 8A 到 8D 是用于描

述逻辑元素的自动确定的图。图 9A 到 9B 是用于描述逻辑结构模型存储部件（转换后）的图。图 10A 和 10B 是用于描述纠错操作的图。

[0051] 参照图 2，根据第一实施方式的图像识别装置 10 包括图像输入单元 11、用户操作输入单元 12、显示单元 13、第一存储单元 20、第二存储单元 25 以及控制单元 30。

[0052] 图像输入单元 11 接收表格图像。具体的是，图像输入单元 11 接收要通过图像识别装置 10 识别的表格图像，并将该表格图像提供给下面描述的字符串识别器 31。例如，可以通过扫描仪实现图像输入单元 11，或者可以通过将另一装置读取的表格图像（图像数据）预先存储在存储器中并将该表格图像提供给字符串识别器 31 的功能来实现图像输入单元 11。

[0053] 现在将利用实施例来描述图像输入单元 11 接收到的表格图像。例如，图像输入单元 11 接收图 3 中示出的表格图像。图 3 中示出的表格图像具有这样的逻辑结构，其中“姓名”、“地址”以及“电话号码”被定位于“用户信息”的下方，“地址”和“电话号码”被定位于“办公信息”的下方，而“用户信息”被定位在与“办公信息”相同的层次内。图 3 中示出的表格图像的逻辑结构指示这样的位置关系，其中在“用户信息”中的“姓名”下方描述“办公信息”中的“地址”和“电话号码”，在所述“用户信息”下方描述用户信息中的“地址”和“电话号码”。图像识别装置 10 要识别的图像并不限于图 3 中示出的表格图像，而可以是任何图像，例如其中为多个问题给出答案的问卷图像，只要字符串被表述在每个项中即可。图像识别装置 10 要识别的图像并不限于其中手写字符串的图像，例如图 3 中示出的图像。

[0054] 回来参照图 2，用户操作输入单元 12 接收用户确认图像识别装置 10 的识别结果的操作。具体的是，操作输入单元 12 接收用于确定多个逻辑元素中的特定逻辑元素为正确地确定请求的输入，并且将该确定请求提供给下面描述的第一逻辑元素确定器 34。用户操作输入单元 12 通常由鼠标和键盘实现。

[0055] 显示单元 13 显示图像识别装置 10 中的识别结果。具体的是，显示单元 13 显示由下面描述的识别结果显示缓冲器 26 中存储的、通过下面描述的识别结果输出器 33 输出的识别结果。显示单元 13 通常通过显示设备实现。注意，显示单元 13 对应于权利要求书中描述的“输出器”。

[0056] 现在将利用实施例来描述显示单元 13 中显示的识别结果。例如，显示单元 13 将图 4 中右半部分中示出的屏面显示为识别结果显示屏面。在图 4 中的左半部分中示出要识别的表格图像，而在图 4 中的右半部分中示出识别结果。在本实施例中，通过比较图 4 的左半部分和右半部分，显然图像识别装置 10 错误地颠倒了用户信息中的“地址”和“电话号码”与办公信息中的“地址”和“电话号码”。

[0057] 返回参照图 2，诸如只读存储器 (ROM) 的第一存储单元 20 存储控制单元 30 执行的各种处理中使用的数据。第一存储单元 20 包括作为与第一实施方式紧密相关的组件的字符识别字典存储部件 21 和逻辑结构模型存储部件 22。

[0058] 字符识别字典存储部件 21 存储在通过使用诸如 OCR 的图像识别技术的字符串识别中使用的字典。具体的是，字符识别字典存储部件 21 存储字符串的识别中使用的字典，并且存储的字典被用于由下面描述的字符串识别器 31 进行的处理。

[0059] 逻辑结构模型存储部件 22 存储由图像识别装置 10 用来识别表格图像的逻辑结构的逻辑结构模型。具体的是，逻辑结构模型存储部件 22 存储用于表格图像中的每个逻辑元

素的识别的逻辑结构模型。存储的逻辑结构模型被用于由下面描述的逻辑结构识别器 32 进行的处理。例如，逻辑结构模型存储部件 22 存储这样的逻辑结构模型，其中的逻辑结构与逻辑元素由用户等使用图像识别装置 10 预先适当设计并准备。

[0060] 现在将利用实施例来描述逻辑结构模型存储部件 22 中存储的逻辑结构模型。例如，逻辑结构模型存储部件 22 存储图 5A 和 5B 中示出的逻辑结构模型。实际上，图 5A 和 5B 中示出的逻辑结构模型在概念上可由实际上存储在第一存储单元 20 中的如图 5A 中示出的树结构和如图 5B 中示出的表结构来表示。

[0061] 在图 5B 中示出的表结构中，在每行中描述每个逻辑元素的类型。从左开始按顺序描述每个逻辑元素的 ID 号、逻辑元素的名称、树结构中逻辑元素的层次、树结构中逻辑元素的父节点的 ID 号、逻辑元素之间的位置关系、属性以及数据字符串的内容。例如，位置关系指示项“姓”在右侧具有包括“ID ＝ 6”的项“姓名”。在位置关系中描述逻辑元素之间的相对位置关系。树结构中具有实体的项具有“节点”属性，而不具有实体并且为语义相关项的集合的项具有“组”属性。针对每个节点描述字符串应该满足的字符条件。例如，针对“电话号码”定义包括“数字串”和“-”的字符类型。

[0062] 返回参照图 2，诸如随机访问存储器 (RAM) 的第二存储单元 25 存储在由控制单元 30 执行的各种处理中使用的数据。如图 2 中所示，第二存储单元 25 包括作为与第一实施方式紧密相关的组件的识别结果显示缓冲器 26。

[0063] 识别结果显示缓冲器 26 存储图像识别装置 10 的识别结果。具体的是，识别结果显示缓冲器 26 存储下面描述的字符串识别器 31 和逻辑结构识别器 32 的识别结果。存储的识别结果被用于下面描述的识别结果输出器 33 中的处理，并且被显示在显示单元 13 中。

[0064] 现在将利用实施例来描述识别结果显示缓冲器 26 中存储的识别结果。例如，识别结果显示缓冲器 26 存储图 6 中示出的识别结果。在图 6 中示出，根据对不同类型逻辑元素的字符串的识别而产生的识别结果（包括候选识别结果），和指示每个字符串被描述为哪种逻辑元素的识别结果（包括候选识别结果）。

[0065] “用户信息”中“姓名”中的逻辑元素“姓”被识别为字符串“Smith”，并且不存在其他候选。“用户信息”中“姓名”中的逻辑元素“名”被识别为第一候选字符串“David”和第二候选字符串“Devid”。针对“用户信息”中的逻辑元素“地址”识别出第一到第四候选字符串。候选中“777”和“111”的识别指示出字符串的识别产生多个候选。候选中“11th AvenueEast，Suite11，Seattle，WA”和“Apple Street，#111，Seattle，WA”的识别指示每个字符串被描述为其逻辑元素的识别产生多个候选。相同的识别操作也应用于其他逻辑元素。

[0066] 返回参照图 2，控制单元 30 控制图像识别装置 10 以执行各种处理。控制单元 30 包括作为与第一实施方式紧密相关的组件的字符串识别器 31、逻辑结构识别器 32、识别结果输出器 33、第一逻辑元素确定器 34、第二逻辑元素确定器 35、逻辑结构重新识别控制器 36 以及逻辑结构模型转换器 37。识别结果输出器 33 对应于权利要求中描述的“输出装置”；第一逻辑元素确定器 34 对应于其中描述的“第一确定装置”；第二逻辑元素确定器 35 对应于其中描述的“第二确定装置”；并且逻辑结构重新识别控制器 36 和逻辑结构模型转换器 37 对应于其中描述的“重新识别装置”。

[0067] 在控制单元 30 中，字符串识别器 31 通过诸如 OCR 的图像识别技术来识别字符串。

具体的是，字符串识别器 31 通过使用存储在字符识别字典存储部件 21 中的字典来识别字符串的字符串和候选（多个候选）在由图像输入单元 11 接收到的表格图像中的位置。识别结果被用于由逻辑结构识别器 32 进行的处理。

[0068] 如上面所描述的，在图 6 中示出的识别结果显示缓冲器的描述中，例如，表格图像中字符串“David”的作为第一候选“David”和第二候选“Devid”的识别由字符串识别器 31 执行。

[0069] 逻辑结构识别器 32 识别每个字符串被描述为哪种逻辑元素。具体的是，逻辑结构识别器 32 参照逻辑结构模型存储部件 22 中存储的逻辑结构模型，以将逻辑结构模型的内容与由字符串识别器 31 识别的识别结果相关联。这些识别结果被存储在识别结果显示缓冲器 26 中，并且被用于由识别结果输出器 33 进行的处理。

[0070] 如上面所述，在图 6 中示出的识别结果显示缓冲器的描述中，例如，“用户信息”中逻辑元素“地址”的作为第一到第四候选的“77711thAvenue East，Suite11，Seattle，WA”、“11111th Avenue East，Suite11，Seattle，WA”、“777Apple Street，#111，Seattle，WA”，以及“111Apple Street，#111，Seattle，WA”的识别由逻辑结构识别器 32 执行。

[0071] 识别结果输出器 33 以特定顺序输出作为识别结果的多个逻辑元素。具体的是，当作为存储在识别结果显示缓冲器 26 中的识别结果的逻辑元素被字符串识别器 31 和逻辑结构识别器 32 识别或重新识别时，识别结果输出器 33 以特定顺序在显示单元 13 中显示逻辑元素。

[0072] 如上面所述，在显示单元 13 的描述中，识别结果输出器 33 输出例如图 4 中示出的识别结果显示屏面。根据第一实施方式的识别结果输出器 33 垂直排列逻辑元素，来显示这些逻辑元素。

[0073] 当接收到用于确定特定逻辑元素为正确的确定请求的输入时，第一逻辑元素确定器 34 确定多个逻辑元素中的特定逻辑元素为正确。具体的是，当第一逻辑元素确定器 34 接收到用户通过用户操作输入单元 12 输入的确定请求的输入时，第一逻辑元素确定器 34 确定接收到的逻辑元素为正确，并向第二逻辑元素确定器 35 指示被确定正确的逻辑元素，所述用户对通过识别结果输出器 33 显示在显示单元 13 中的且被存储在识别结果显示缓冲器 26 中的识别结果进行确认。

[0074] 现在将利用实施例来描述第一逻辑元素确定器 34 进行的确定。例如，假设图 7 中示出的识别结果显示屏面通过识别结果输出器 33 显示在显示单元 13 中。还假设已经确认识别结果的用户确定“用户信息”中“地址”中的字符串“77711th Avenue East，Suite11，Seattle，WA”是错误项，并且希望将该项纠正为“111Apple Street，#111，Seattle，WA”。当用户使用用户操作输入单元 12（在本实施例中为鼠标）将指针移动到“77711th AvenueEast，Suite11，Seattle，WA”，并右键点击错误项（参照图 7 中的 (1)）以纠正该错误项时，随后的候选字符串“11111th Avenue East，Suite11，Seattle，WA”、“777Apple Street，#111，Seattle，WA”以及“111Apple Street，#111，Seattle，WA”显示在窗口中（参照图 7 中的 (2)）。这些随后的候选是存储在识别结果显示缓冲器 26 中的识别结果（参照图 6）。

[0075] 接着，当用户使用用户操作输入单元 12（鼠标）左键点击“111AppleStreet，#111，Seattle，WA”时，“111Apple Street，#111，Seattle，WA”被输入为针对“地址”的字符串

（图 7 中未示出），并且例如，“确定”图标改变为词语“已确定”。

[0076] 第一逻辑元素确定器 34 在用户点击“确定”图标以清楚地指令进行指定项的确定时确定所述指定项为正确，而不是在用户纠正错误项时确定纠正后的项为正确。

[0077] 返回参照图 2，第二逻辑元素确定器 35 确定在特定逻辑元素之前输出的所有逻辑元素的正确性。具体的是，当第一逻辑元素确定器 34 确定特定逻辑元素为正确时，第二逻辑元素确定器 35 确定在该特定逻辑元素之前输出并且根据用户进行的确认定位的所有逻辑元素的正确性，并向逻辑结构重新识别控制器 36 指示被确定为正确的逻辑元素。

[0078] 假设用户通常从顶至底按顺序逐一确认图 7 中示出的识别结果显示屏面中的项。在这种情况中，在确定的逻辑元素（用户信息中的“地址”）之前被输出的、并且根据用户的确认被定位的逻辑元素指的是定位在用户信息中的“地址”上方的逻辑元素。用户信息中的“姓名”对应于在确定的逻辑元素之前输出并根据用户的确认被定位的逻辑元素。因此，第二逻辑元素确定器 35 还确定用户信息中的“姓”和“名”正确。换言之，第二逻辑元素确定器 35 例如通过对逻辑元素在表（如图 6 中示出的）中被描述在哪行进行计数以知道行号，来获取存储在识别结果显示缓冲器 26 中的显示顺序，并确定以较早顺序（例如具有小行号）显示的所有逻辑元素的正确性。

[0079] 具体的是，第二逻辑元素确定器 35 将存储在识别结果显示缓冲器 26 内的识别结果中由第一逻辑元素确定器 34 确定的“用户信息”中的逻辑元素“地址”确定为“111Apple Street，#111，Seattle，WA”，并且例如，向第二及其后的候选指派禁用标记，以实际删除这些候选。然后，第二逻辑元素确定器 35 将“用户信息”中的逻辑元素“姓”确定为“Smith”，并将逻辑元素“名”确定为“David”，并且向第二候选“Devid”指派禁用标记，以实际删除第二候选。例如图 8A 中示出的信息被存储其中反映确定的内容的识别结果显示缓冲器 26 中（有阴影线的逻辑元素指示已确定的逻辑元素）。

[0080] 逻辑结构重新识别控制器 36 基于逻辑元素的确定内容来控制逻辑结构识别器 32 和逻辑结构模型转换器 37，以重新识别仍未被确定为正确的逻辑元素的对应关系。具体的是，逻辑结构重新识别控制器 36 基于通过第一逻辑元素确定器 34 和第二逻辑元素确定器 35 确定为正确的逻辑元素的确定内容，来控制逻辑结构识别器 32 和逻辑结构模型转换器 37，以促使逻辑结构识别器 32 重新识别仍未被确定为正确的逻辑元素与字符串之间的对应关系。

[0081] 逻辑结构模型转换器 37 转换逻辑结构模型。具体的是，逻辑结构模型转换器 37 受逻辑结构重新识别控制器 36 的控制，并且反映了存储在逻辑结构模型存储部件 22 中的逻辑结构模型中的通过第一逻辑元素确定器 34 和第二逻辑元素确定器 35 获得的确定内容。逻辑结构模型中的反映指的是将确定内容描述在逻辑结构模型中，以唯一地确定对应的字符串。

[0082] 现在将利用实施例来描述通过逻辑结构模型转换器 37 转换的逻辑结构模型。例如，逻辑结构模型存储部件 22 存储图 9A 中示出的根据逻辑结构模型转换器 37 的转换而产生的逻辑结构模型。如从图 5B 和图 9A 之间的比较所见，逻辑结构模型转换器 37 分别描述了“用户信息”中“地址”、“姓”以及“名”内的“111Apple Street，#111，Seattle，WA”、“Smith”以及“David”，以唯一地确定对应字符串。创建其中描述了通过逻辑结构模型转换器 37 的转换而获得的已确定内容的逻辑结构模型。

[0083] 返回参照逻辑结构重新识别控制器 36 的描述，在促使逻辑结构模型转换器 37 转换存储在逻辑结构模型存储部件 22 中的逻辑结构模型之后，逻辑结构重新识别控制器 36 控制逻辑结构识别器 32，以执行重新识别。逻辑结构识别器 32 参照经受转换并存储在逻辑结构模型存储部件 22 中的逻辑结构模型，并将该逻辑结构模型的内容与字符串识别器 31 识别出的识别结果相关联。这些识别结果同样被存储在识别结果显示缓冲器 26 中，并被用于通过识别结果输出器 33 进行的处理。

[0084] 具体的是，假设获得了通过逻辑结构识别器 32 进行的识别的结果（即，存储在图 8B 中示出的识别结果显示缓冲器 26 中的识别结果）。换言之，假设“用户信息”中“地址”的确定将“用户信息”中的“电话号码”限制为以“999”开始的号码、将“办公信息”中的“地址”限制为包括“11th Avenue East，Suite11”的地址，并将“办公信息”中的“电话号码”限制为以“888”开始的号码。

[0085] 在上面的情况中，识别结果输出器 33 在显示单元 13 中显示图 10A 中示出的识别结果显示屏面。在图 10A 中示出的识别结果显示屏面中，注意“用户信息”中“姓名”和“地址”右侧的“确定”图标被改变为词语“已确定”，意思是字符串已被确定。假设已经确认了识别结果的用户确定“用户信息”中“电话号码”中的“999-666-7777”为错误项，并且此时希望将该项纠正为“999-666-1111”。当用户使用用户操作输入单元 12（鼠标）将指针移动到“999-666-7777”上，并右键点击该错误项以纠正该错误项时，随后的候选字符串“999-666-1111”被显示在窗口中。接着，当用户使用用户操作输入单元 12（鼠标）左键点击“999-666-1111”时，“999-666-1111”被输入为针对“电话号码”的字符串，并且“用户信息”中“电话号码”右侧的“确定”图标被改变为词语“已确定”，如图 10B 中所示出。

[0086] 此时，第二逻辑元素确定器 35 将存储在识别结果显示缓冲器 26 内的识别结果中的由第一逻辑元素确定器 34 确定的“用户信息”中的逻辑元素“电话号码”确定为“999-666-1111”，并且向第二候选指派禁用标记，以实际删除该候选（参照图 8C）。

[0087] 然后，逻辑结构重新识别控制器 36 促使逻辑结构模型转换器 37 转换存储在逻辑结构模型存储部件 22 中的逻辑结构模型（参照图 9B），并控制逻辑结构识别器 32，以执行重新识别。然而，在这种情况中，识别结果未被改变（参照图 8D），并且识别结果输出器 33 保持图 10B 中示出的识别结果显示屏面的显示。

[0088] 根据第一实施方式的图像识别装置执行的处理

[0089] 现在将参照图 11 描述根据第一实施方式的图像识别装置执行的处理。图 11 是示出根据第一实施方式的图像识别装置执行的处理的实施例的流程图。

[0090] 参照图 11，在步骤 S101 中，图像识别装置 10 中的字符串识别器 31 确定是否从图像输入单元 11 接收到表格图像的输入。如果接收到表格图像的输入（步骤 S101 中的确定为肯定的），则在步骤 S102 中，字符串识别器 31 识别字符串并将识别结果提供给逻辑结构识别器 32。

[0091] 在步骤 S103 中，逻辑结构识别器 32 参照存储在逻辑结构模型存储部件 22 中的逻辑结构模型，并将逻辑结构模型的内容与字符串识别器 31 识别的识别结果相关联。逻辑结构识别器 32 将识别结果提供给识别结果输出器 33。

[0092] 在步骤 S104 处，识别结果输出器 33 以特定顺序在显示单元 13 中显示作为识别结果的逻辑元素。

[0093] 在步骤 S105 中，第一逻辑元素确定器 34 确定是否按下所有项确定按钮，例如图 4 中示出的识别结果显示屏面的右半部分中底部处的“确定”图标。如果按下了所有项确定按钮（步骤 S105 中的确定为肯定的），则处理终止。

[0094] 如果未按下所有项确定按钮（步骤 S105 中的确定为否定的），则在步骤 S106 中，第一逻辑元素确定器 34 确定是否执行纠错操作。如果执行纠错操作（步骤 S106 中的确定为肯定），则在步骤 S107 中，第一逻辑元素确定器 34 确定针对其执行纠错操作的特定逻辑元素的正确性，并向第二逻辑元素确定器 35 指示被确定为正确的逻辑元素。如果纠错操作未被执行（步骤 S106 中的确定为否定），则在步骤 S108 中，第一逻辑元素确定器 34 确定是否按下指定项确定按钮，例如图 4 中示出的识别结果显示屏面的右半部分中每个项右侧的“确定”图标。如果按下指定项确定按钮（步骤 S108 中的确定为肯定），则在步骤 S109 中，第一逻辑元素确定器 34 确定已确定指定项的逻辑元素为正确，并向第二逻辑元素确定器 35 指示被确定为正确的逻辑元素。

[0095] 当特定逻辑元素被第一逻辑元素确定器 34 确定为正确（步骤 107 和 109）时，第二逻辑元素确定器 35 在步骤 S110 中获取存储在识别结果显示缓冲器 26 中的显示顺序，并获取已确定项的顺序。在步骤 S111 中，第二逻辑元素确定器 35 确定在已确定项的逻辑元素之前输出并且根据用户进行的确认来定位的所有逻辑元素的正确性，并向逻辑结构重新识别控制器 36 指示被确定为正确的逻辑元素。

[0096] 在步骤 S112 中，逻辑结构重新识别控制器 36 确定是否确定了所有项，例如，用户是否按下所有项确定按钮。如果未确定所有项（步骤 S112 中的确定为否定），则逻辑结构重新识别控制器 36 促使逻辑结构模型转换器 37 转换逻辑结构模型存储部件 22 中存储的逻辑结构模型，然后，在步骤 S113 中，逻辑结构重新识别控制器 36 控制逻辑结构识别器 32，以执行重新识别。此时，逻辑结构识别器 32 参照经受转换的逻辑结构模型，以将该逻辑结构模型的内容与字符串识别器 31 识别的识别结果相关联。然后，处理返回到步骤 S104。

[0097] 根据第一实施方式的另一识别结果显示屏面

[0098] 尽管在上面的第一实施方式中描述了图 4 中示出的识别结果显示屏面，但是识别结果显示屏面并不限于图 4 中所示出的识别结果显示屏面。例如，可以采用图 12 中示出的识别结果显示屏面。如上面所描述的，第二逻辑元素确定器 35 确定在特定逻辑元素之前输出并且根据用户进行的确认来定位的所有逻辑元素的正确性。现在将进一步描述“在特定逻辑元素之前根据用户进行的确认定位的顺序”的含义。

[0099] 例如，在图 12 中的下部中示出的识别结果显示屏面中，假设用户通常在顶行以从左到右的顺序逐一确认该行中的项，直到该行的右端处的项。当确认了顶行的右端处的最后项时，用户通常向下到底行，并在该行以从左到右的顺序注逐一确认这些项，直到该行的右端处的项。因为这里的“确认”并未伴随有用户进行的任何操作，所以该确认的顺序可以是任意的。然而，可以根据画面的操作模式或设计预先假定确认的顺序。根据这种观点，根据第一实施方式的图像识别装置假定“在特定逻辑元素之前根据用户进行的确认定位的顺序”，并且例如，当识别结果被存储在通过用户使用图像识别装置预先设置的识别结果显示缓冲器中时，设置 ID 的升序。此外，根据这种观点，在用户确定特定逻辑元素之后，在该特定逻辑元素之前输出并且根据用户进行的确认定位的所有逻辑元素可以被认为是已确定。

[0100] 在图 12 中示出的识别结果显示屏面中，当已确定逻辑元素为办公“地址”时，第二

逻辑元素确定器 35 认为用户信息中的“姓名”、“地址”以及“电话号码”是在该办公“地址”之前输出并根据用户进行的确认定位的逻辑元素，并确定这些逻辑元素正确。

[0101] 如上面所描述的，根据本发明的第一实施方式，能够减小纠正操作的负荷。具体的是，根据第一实施方式的图像识别装置参照具有通过选定用户确定或纠正的逻辑元素的逻辑结构模型，以改善其他未确定逻辑元素的识别结果。如果针对未确定逻辑元素的识别结果的候选数被降低为 1，则能够确定该逻辑元素。因此，由于用户确定或纠正的结果可以被用来确定其他逻辑元素，因此能够减少用户进行纠正操作的负荷。

[0102] 此外，根据本发明的第一实施方式，在确定特定逻辑元素之后，自动确定位于已确定逻辑元素之前的逻辑元素。因此，能够自动确定大量未确定的逻辑元素，由此进一步减小了用户进行纠正操作的负荷。

[0103] 此外，本发明的第一实施方式在识别具有未知格式的表格图像以记录每个项的字符串数据时非常有效，并且有助于改善使用表格图像进行的数据输入操作的效率。因此，能够便于在纸质表格与计算机化数据之间进行交换，并改善使用纸质表格进行操作的效率。

[0104] 第二实施方式

[0105] 在本发明的第一实施方式中描述了，当图像识别装置重新识别表格图像时，立即在识别结果显示屏面中反映重新识别结果的方法。然而，替代立即在识别结果显示屏面中反映重新识别结果的方法，可以采用显示如下信息的显示方法，所述信息指示出重新识别结果与当前在输出单元中显示的识别结果不同。这种方法被描述为本发明的第二实施方式。

[0106] 根据第二实施方式的图像识别装置的概况

[0107] 现在将参照图 13A 和 13B 描述根据第二实施方式的图像识别装置的概况和特征。图 13A 和 13B 是描述根据第二实施方式的图像识别装置的概况的图。

[0108] 通过该图像识别装置进行的表格图像的重新识别可以改善未确定逻辑元素的识别结果。换言之，从用户观点来看，一个项的纠正会自动纠正其他项的识别结果。例如，假设显示的多个逻辑元素包括一个错误项。在这种情况中，自然考虑通过用户对一个错误项进行纠正而使所有识别结果都正确。然而，当在用户纠正错误项之后执行重新识别时，正确显示的未确定正确项可能被变为错误项。尽管重新识别通常可以改善识别结果，但是不能否认使正确结果变坏的可能性。换言之，未确定逻辑元素的自动改变可能减小针对用户的可用性。

[0109] 为了解决上面的问题，如图 13A 和 13B 中所示出的，当在错误项被纠正之后执行的重新识别导致针对未确定逻辑元素的第一候选的识别结果的改变时，根据第二实施方式的图像识别装置并不简单地改变该项的字符串，而是高亮显示被改变的识别结果。

[0110] 具体的是，当在用户信息中的错误项“地址”如图 13A 中所示出，从“77711th Avenue East，Suite11，Seattle，WA”改变为“111Apple Street，#111，Seattle，WA”之后执行重新识别，以改变针对用户信息中“电话号码”（所述“电话号码”为从“888-444-5555”到“999-666-7777”的未确定逻辑元素）的第一候选的识别结果时，如图 13B 中所示出，图像识别装置高亮显示被改变的识别结果，而不是简单地改变这些项的字符串。由于使用户充分关注对高亮显示项的确认，所以可以期待减小操作负荷。

[0111] 根据第二实施方式的图像识别装置的结构

[0112] 现在将参照图 14 描述根据第二实施方式的图像识别装置的结构。图 14 是示出根据第二实施方式的图像识别装置的结构的实施例的框图。

[0113] 在图 14 中示出的框图中，用粗线环绕的组件被添加到根据第一实施方式的图像识别装置的组件中。具体的是，除根据第一实施方式的图像识别装置中的组件之外，根据第二实施方式的图像识别装置 10 还包括控制单元 30 中的重新识别结果缓冲器存储部件 38 和第二存储单元 25 中的重新识别结果缓冲器 27。

[0114] 重新识别结果缓冲器存储部件 38 存储重新识别结果缓冲器 27 中的重新识别结果。重新识别结果缓冲器 27 是具有与重新识别结果显示缓冲器 26 类似的结构的存储器区。例如，图 15B 中示出的信息被存储在重新识别结果缓冲器 27 中。

[0115] 当输出重新识别的结果时，识别结果输出器 33 针对每个逻辑元素，将通过重新识别结果缓冲器存储部件 38 存储在重新识别结果缓冲器 27 中的重新识别结果（参照图 15B）与当前在显示单元 13 中显示的识别结果（参照图 15A）进行比较。如果存储在重新识别结果缓冲器 27 中的逻辑元素的重新识别结果与当前在显示单元 13 中显示的识别结果不一致，则识别结果输出器 33 高亮显示存储在重新识别结果缓冲器 27 中的重新识别结果，以指示重新识别的结果与当前在显示单元 13 中显示的识别结果不同。

[0116] 根据第二实施方式的图像识别装置执行的处理

[0117] 现在将参照图 16 描述根据第二实施方式的图像识别装置执行的处理。图 16 是示出根据第二实施方式的图像识别装置执行的处理的实施例的流程图。

[0118] 如从图 11 和图 16 之间的比较所见，根据第二实施方式的图像识别装置执行的处理与根据第一实施方式的图像识别装置执行的处理的不同之处仅在于增加了步骤 S214。

[0119] 具体的是，在根据第二实施方式的图像识别装置中，在步骤 S212 中，逻辑结构重新识别控制器 36 确定是否确定了所有项。如果未确定所有项（步骤 S212 中的确定为否定），则逻辑结构重新识别控制器 36 促使逻辑结构模型转换器 37 转换逻辑结构模型存储部件 22 中存储的逻辑结构模型，然后，在步骤 S213 中，逻辑结构重新识别控制器 36 控制逻辑结构识别器 32，以执行重新识别。在步骤 S214 中，重新识别结果缓冲器存储部件 38 将重新识别的结果存储在重新识别结果缓冲器 27 中。

[0120] 然后，在步骤 S204 中，识别结果输出器 33 针对每个逻辑元素，将存储在重新识别结果缓冲器 27 中的重新识别结果与存储在识别结果显示缓冲器 26 中的识别结果进行比较。如果存储在重新识别结果缓冲器 27 中的逻辑元素的重新识别结果与存储在识别结果显示缓冲器 26 中的识别结果不一致，则识别结果输出器 33 高亮显示存储在重新识别结果缓冲器 27 中的重新识别结果，以指示该重新识别的结果与当前在显示单元 13 中显示的识别结果不同。

[0121] 根据第二实施方式的另一识别结果显示屏面

[0122] 尽管上面参照图 13A 和 13B 描述的识别结果显示屏面被假设为第二实施方式中的识别结果显示屏面，但是识别结果显示屏面并不限于上面参照图 13A 和 13B 描述的识别结果显示屏面。例如，可以采用参照图 17A 和 17B 描述的识别结果显示屏面。具体的是，如图 17A 和 17B 中所示出，图像识别装置反映针对如下字符串的候选中的重新识别结果，即当即使在纠正错误项之后也在不改变未确定逻辑元素的显示的情况下选择未确定逻辑元素时显示的所述字符串。

[0123] 例如，当在用户信息中的错误项“地址”从“77711th Avenue East，Suite11，Seattle，WA”改变为如图 17A 中所示出的“111Apple Street，#111，Seattle，WA”之后执行重新识别，以将针对用户信息中的未确定逻辑元素“电话号码”的第一候选的识别结果从“888-444-5555”改变为“999-666-7777”时，图像识别装置并不立即改变用户信息中“电话号码”的输出。图像识别装置保持“888-444-5555”的显示，并反映针对在选择用户信息中的未确定逻辑元素“电话号码”时显示的字符串的候选中的重新识别的结果（“999-666-7777”）（参照图 17B）。

[0124] 在这种情况中，即使在执行重新识别之后，重新识别的结果也不被立即存储到识别结果显示缓冲器 26 中，并且不显示在显示单元 13 中；换言之，重新识别的结果未被呈现给用户。当用户选择一项以纠正该项的识别结果时，该项的纠正被指示给识别结果输出器 33，在所述识别结果输出器 33 中仅将用户选择的项从重新识别结果缓冲器 27 复制到识别结果显示缓冲器 26。按照上面的方式，仅呈现用户选择要进行纠错的项的重新识别的结果。

[0125] 如上面所述，根据本发明的第二实施方式，能够防止未确定逻辑元素（未被确定为正确的逻辑元素）变得更差。

[0126] 第三实施方式

[0127] 在第二实施方式中描述了一种方法，其中当图像识别装置从用户接收到用于确定特定逻辑元素的确定请求的输入时，图像识别装置确定在特定逻辑元素之前输出的所有逻辑元素为正确的，以执行重新识别并将重新识别的结果临时存储在缓冲器中。然而，“将重新识别的结果临时存储在缓冲器中”的方法也可应用于其中图像识别装置在分别确定逻辑元素时执行重新识别的情况。

[0128] 换言之，只要使用“将重新识别的结果临时存储在缓冲器中”的方法就具有如下优点：减小了用户进行纠正操作的负荷，并且防止了未确定逻辑元素变得更差。因此，不论图像识别装置是分别确定逻辑元素，还是自动确定位于特定逻辑元素之前的所有逻辑元素，都可以采用这种方法。

[0129] 根据第三实施方式的图像识别装置的概况

[0130] 现在将参照图 18A 和 18B 以及 19A 和 19B 来描述根据本发明第三实施方式的图像识别装置的概况和特征。图 18A 和 18B 是用于描述根据第三实施方式的图像识别装置的概况和特征的图。图 19A 和 19B 是用于描述识别结果显示缓冲器和重新识别结果缓冲器的图。

[0131] 例如，根据第三实施方式的图像识别装置在显示单元 13 中显示图 18A 和 18B 中示出的识别结果显示屏面。在图 18A 中示出的识别结果显示屏面中，假设已经确认识别结果的用户确定针对用户信息中的“地址”的字符串“77711th Avenue East，Suite11，Seattle，WA”为错误项，并将该项纠正为“111Apple Street，#111，Seattle，WA”。图 18B 示出根据该纠正产生的识别结果显示屏面。注意，“确定”图标被保持显示在“用户信息”中的“姓名”的右侧并且未被改变为词语“已确定”。

[0132] 换言之，即使当根据第三实施方式的图像识别装置接收用于确定特定逻辑元素为正确（用户信息中的“地址”）的确定请求的输入时，该图像识别装置也并不确定在特定逻辑元素之前根据通过用户的确认的定位顺序输出的逻辑元素（用户信息中的“姓名”）的正确性，而是，该图像识别装置将该逻辑元素作为未确定的逻辑元素来处理。

[0133] 还要注意，在图 18B 中高亮显示用户信息中的“电话号码”和办公信息中的“地址”

和“电话号码”。

[0134] 当用户信息中的“地址”从“77711th Avenue East,Suite11,Seattle,WA”改变为“111Apple Street,#111,Seattle,WA”之后，执行重新识别时，根据第三实施方式的图像识别装置并不高亮显示其识别结果未被改变的未确定逻辑元素（用户信息中的“姓名”）。替代的是，根据第三实施方式的图像识别装置高亮显示其识别结果被改变的未确定逻辑元素（用户信息中的“电话号码”和办公信息中的“地址”和“电话号码”）。

[0135] 现在将参照图 19A 和 19B 来描述此时的识别结果显示缓冲器 26 和重新识别结果缓冲器 27。图 19A 示出用户已经确定用户信息中的“地址”之后，用户信息中的“地址”被重新识别之前的识别结果显示缓冲器 26。在图 19A 中示出的识别结果显示缓冲器 26 中，用户信息中的“姓”和“名”并未被识别为已确定逻辑元素。图 19B 示出用户已经确定用户信息中的“地址”并且已经执行重新识别之后的重新识别结果缓冲器 27。在图 19B 中示出的重新识别结果缓冲器 27 中，仅高亮显示其第一候选与图 19A 中的识别结果显示缓冲器 26 中的第一候选不同的逻辑元素（仅用户信息中的“电话号码”和办公信息中的“地址”与“电话号码”）。

[0136] 根据第三实施方式的图像识别装置执行的处理

[0137] 现在将参照图 20 来描述根据第三实施方式的图像识别装置执行的处理。图 20 是示出根据第三实施方式的图像识别装置执行的处理的实施例的流程图。

[0138] 如从图 16 和图 20 之间的比较所见，根据第三实施方式的图像识别装置执行的处理与根据第二实施方式的图像识别装置执行的处理的不同之处仅在于，因为不需要第二逻辑元素确定器 35，所以删除了图 16 中的步骤 S210 和 S211。

[0139] 具体的是，在根据第三实施方式的图像识别装置中，在步骤 S305 中，第一逻辑元素确定器 34 确定是否按下所有项确定按钮。如果未按下所有项确定按钮（步骤 S305 中的确定为否定），则在步骤 S306 中，第一逻辑元素确定器 34 确定是否执行纠错操作。如果执行纠错操作（步骤 S306 中的确定为肯定），则在步骤 S307 中，第一逻辑元素确定器 34 确定针对其执行纠错操作的特定逻辑元素的正确性，并向逻辑结构重新识别控制器 36 指示被确定为正确的逻辑元素。

[0140] 如果不执行纠错操作（步骤 S306 中的确定为否定），则在步骤 S308 中，第一逻辑元素确定器 34 确定是否按下指定项确定按钮。如果按下了指定项确定按钮（步骤 S308 中的确定为肯定），则在步骤 S309 中，第一逻辑元素确定器 34 确定已确定指定项的逻辑元素为正确，并向逻辑结构重新识别控制器 36 指示被确定为正确的逻辑元素。

[0141] 在步骤 S310 中，逻辑结构重新识别控制器 36 确定是否确定了所有项。如果未确定所有项（步骤 S310 中的确定为否定），则逻辑结构重新识别控制器 36 促使逻辑结构模型转换器 37 转换逻辑结构模型存储部件 22 中存储的逻辑结构模型，然后，在步骤 S311 中，逻辑结构重新识别控制器 36 控制逻辑结构识别器 32，以执行重新识别。在步骤 S312 中，重新识别结果缓冲器存储部件 38 将重新识别的结果存储在重新识别结果缓冲器 27 中。

[0142] 然后，在步骤 S304 中，识别结果输出器 33 针对每个逻辑元素，将存储在重新识别结果缓冲器 27 中的重新识别结果与存储在识别结果显示缓冲器 26 中的识别结果进行比较。如果存储在重新识别结果缓冲器 27 中的逻辑元素的重新识别结果与存储在识别结果显示缓冲器 26 中的识别结果不一致，则识别结果输出器 33 高亮显示存储在重新识别结果

缓冲器 27 中的重新识别结果，以指示该重新识别的结果与当前在显示单元 13 中显示的识别结果不同。

[0143] 根据第三实施方式的另一识别结果显示屏面

[0144] 尽管上面参照图 18A 和 18B 描述的识别结果显示屏面被假设为第三实施方式中的识别结果显示屏面，但是识别结果显示屏面并不限于上面参照图 18A 和 18B 描述的识别结果显示屏面。例如，可以采用参照图 21A 和 21B 描述的识别结果显示屏面。具体的是，如图 21A 和 21B 中所示，图像识别装置反映如下字符串的候选中的重新识别结果，即当即使在纠正错误项之后也在不改变未确定逻辑元素的显示的情况下选择未确定逻辑元素时显示的所述字符串。

[0145] 还是在这种情况下，如图 21B 中所示，注意，“确定”图标在“用户信息”中的“姓名”右侧被保持显示，并且未被改变为词语“已确定”。

[0146] 如上面所述，根据本发明的第三实施方式，能够减小纠正操作的负荷。具体的是，根据第三实施方式的图像识别装置参照具有通过选定用户确定或纠正的逻辑元素的逻辑结构模型，以改善其他未确定逻辑元素的识别结果。如果针对未确定逻辑元素的识别结果的候选数被降为 1，则能够确定该逻辑元素。因此，由于用户确定或纠正的结果可以被用来确定其他逻辑元素，所以能够减少用户纠正操作的负荷。

[0147] 此外，本发明的第三实施方式在识别具有未知格式的表格图像，以针对每个项记录字符串数据的情况中非常有效，并且有助于改善使用表格图像进行的数据输入操作的效率。因此，能够便于在纸质表格和计算机化数据之间进行交换，并改善使用纸质表格进行操作的效率。

[0148] 此外，根据本发明的第三实施方式，能够防止未确定的逻辑元素（即未被确定为正确的逻辑元素）变得更差。

[0149] 第四实施方式

[0150] 可以通过不同于上面描述的实施方式的各种实施方式来实施本发明。

[0151] 尽管在上面的实施方式中，图像识别装置包括逻辑结构重新识别控制器，并控制逻辑结构识别器，以在促使逻辑结构模型转换器转换逻辑结构模型之后执行重新识别，但是本发明并不限于上面的方法。可以采用任何基于已确定逻辑元素的确定内容来重新识别未确定逻辑元素的逻辑结构的方法。

[0152] 在上面的实施方式中描述的处理中，可手动执行被描述为自动执行的处理的全部或者一部分，并且可自动执行被描述为手动执行的处理的全部或者一部分。此外，如果未另作说明，则在上述描述中示出的处理（图 11、16 以及 20）、控制处理、指定的姓名、包括各种数据和参数的信息以及附图可以任意变化。

[0153] 在附图中示出的装置中的组件为功能概念，并且可以不必如图（例如图 2 和图 14）中所示地被物理构建。具体的是，每个装置的分布和集成的特定模式并不限于图中示出的模式。根据各种负荷和使用，可以将所有或一些特定模式在功能上或物理上分布或集成在任意单元中。此外，每个装置中执行的处理和功能的全部或者一部分可以通过每个装置的中央处理单元 (CPU) 和由该 CPU 分析并执行的程序来实现，或者可以通过硬布线逻辑来实现。

[0154] 上述实施方式中描述的各种处理可以通过执行预先准备的程序的计算机（例如

个人计算机或工作站）来实现。现在将参照图 22 来描述执行具有类似于根据第一实施方式的功能的图像识别程序的计算机的实施例。图 22 是示出执行图像识别程序的计算机的结构的实施例的框图。

[0155] 参照图 22，图像识别装置（计算机）40 包括通过总线 46 相互连接的高速缓存存储器 41、RAM42、硬盘驱动器 (HDD)、ROM44 以及 CPU45。具有类似于根据第一实施方式的功能的图像识别程序被预先存储在 ROM44 中。具体的是，如图 22 中所示出，字符串识别程序 44a、逻辑结构识别程序 44b、识别结果输出程序 44c、第一逻辑元素确定程序 44d、第二逻辑元素确定程序 44e、逻辑结构重新识别控制程序 44f 以及逻辑结构模型转换程序 44g 被存储在 ROM44 中。

[0156] 如图 22 中所示，CPU45 读取并执行从字符串识别程序 44a 到逻辑结构模型转换程序 44g 的程序，以执行字符串识别处理 45a、逻辑结构识别处理 45b、识别结果输出处理 45c、第一逻辑元素确定处理 45d、第二逻辑元素确定处理 45e、逻辑结构重新识别控制处理 45f 以及逻辑结构模型转换处理 45g。从字符串识别处理 45a 到逻辑结构模型转换程序 44g 的处理分别对应于如图 2 中所示出的字符串识别器 31、逻辑结构识别器 32、识别结果输出器 33、第一逻辑元素确定器 34、第二逻辑元素确定器 35、逻辑结构重新识别控制器 36，以及逻辑结构模型转换器 37。

[0157] 如图 22 中所示，HDD43 包括字符识别字典表 43a 和逻辑结构模型表 43b。字符识别字典表 43a 和逻辑结构模型表 43b 分别对应于图 2 中示出的字符识别字典存储部件 21 和逻辑结构模型存储部件 22。如图 22 中所示，RAM42 包括识别结果显示缓冲器 42a。

[0158] 从字符串识别程序 44a 到逻辑结构模型转换程序 44g 的程序可以不必存储在 ROM44 中。例如，这些程序可以被存储在计算机 40 中装载的“便携式物理介质”（诸如软盘 (FD)、光盘只读存储器 (CD-ROMA)、磁光 (MO) 盘、数字化光盘 (DVD)，或集成电路 (IC) 卡）、“固定的物理介质”（诸如设置在计算机 40 内部或外部的 HDD）、或者通过公共电路、因特网、局域网 (LAN) 或广域网 (WAN) 连接到计算机 40 的“另一计算机（或服务器）”中。在这种情况中，计算机 40 读取并执行来自“便携式物理介质”、“固定的物理介质”或“另一计算机（或服务器）”的程序。

[0159] 本申请基于并要求 2007 年 10 月 31 日递交的在先日本专利申请 NO.2007-284294 的优先权，这里通过引用并入该申请的全部内容。

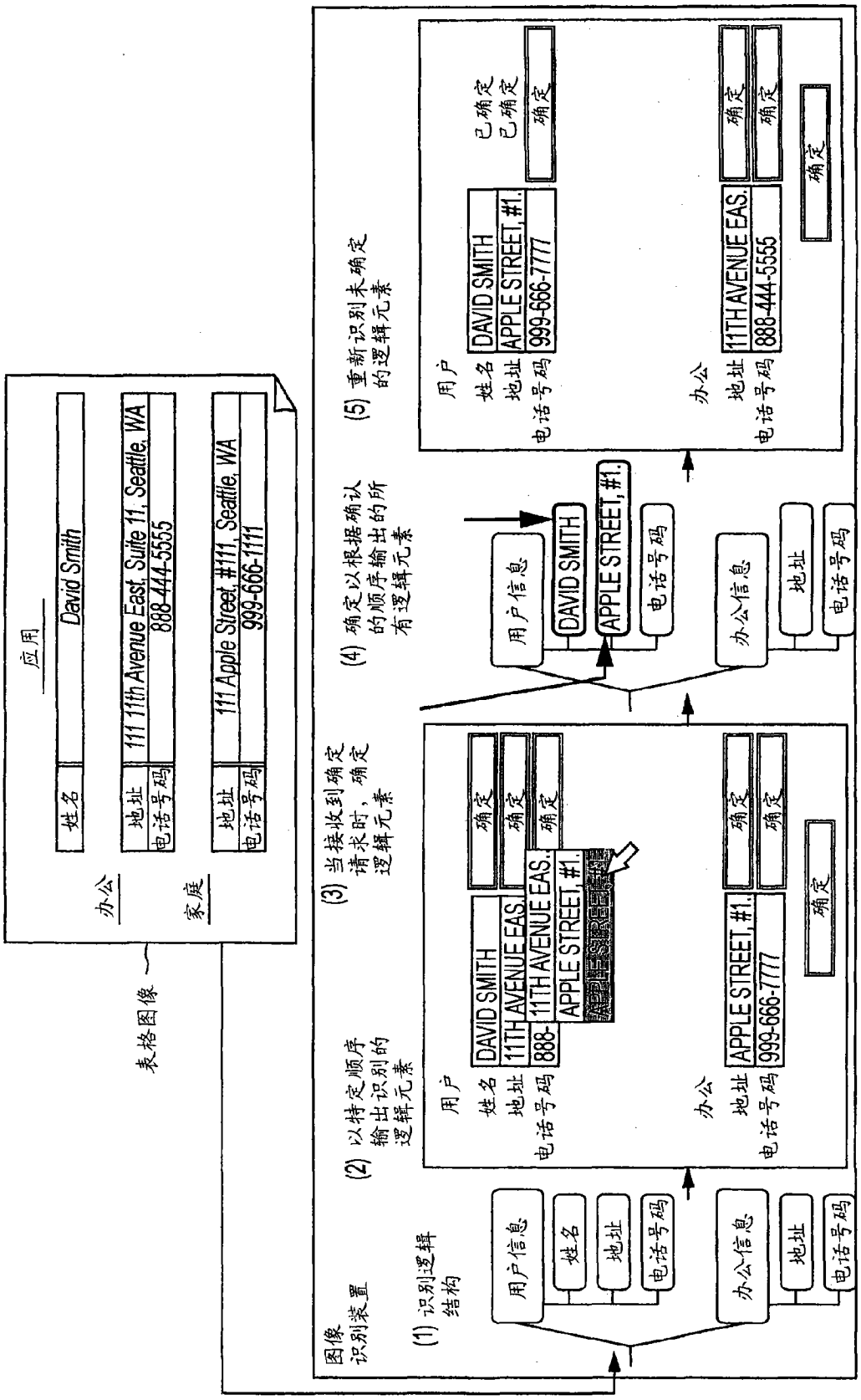

应用
姓名
David Smith
办公
地址
111 11th Avenue East, Suite 11, Seattle, WA
电话号码
888-444-5555
家庭
地址
111 Apple Street, #111, Seattle, WA
电话号码
999-666-1111
表格图像
图像识别装置
(1) 识别逻辑结构
(2) 以特定顺序输出识别的逻辑元素
(3) 当接收到确定请求时，确定逻辑元素
(4) 确定以根据确认的所有逻辑元素的顺序输出
(5) 重新识别未确定的逻辑元素
用户信息
姓名
地址
电话号码
办公信息
地址
电话号码
用户
姓名
地址
电话号码
DAVID SMITH
11TH AVENUE EAS.
888-
APPLE STREET, #1.
确定
办公
地址
电话号码
APPLE STREET, #1.
999-666-7777
确定
已确定
11TH AVENUE EAS.
888-444-5555

图 1

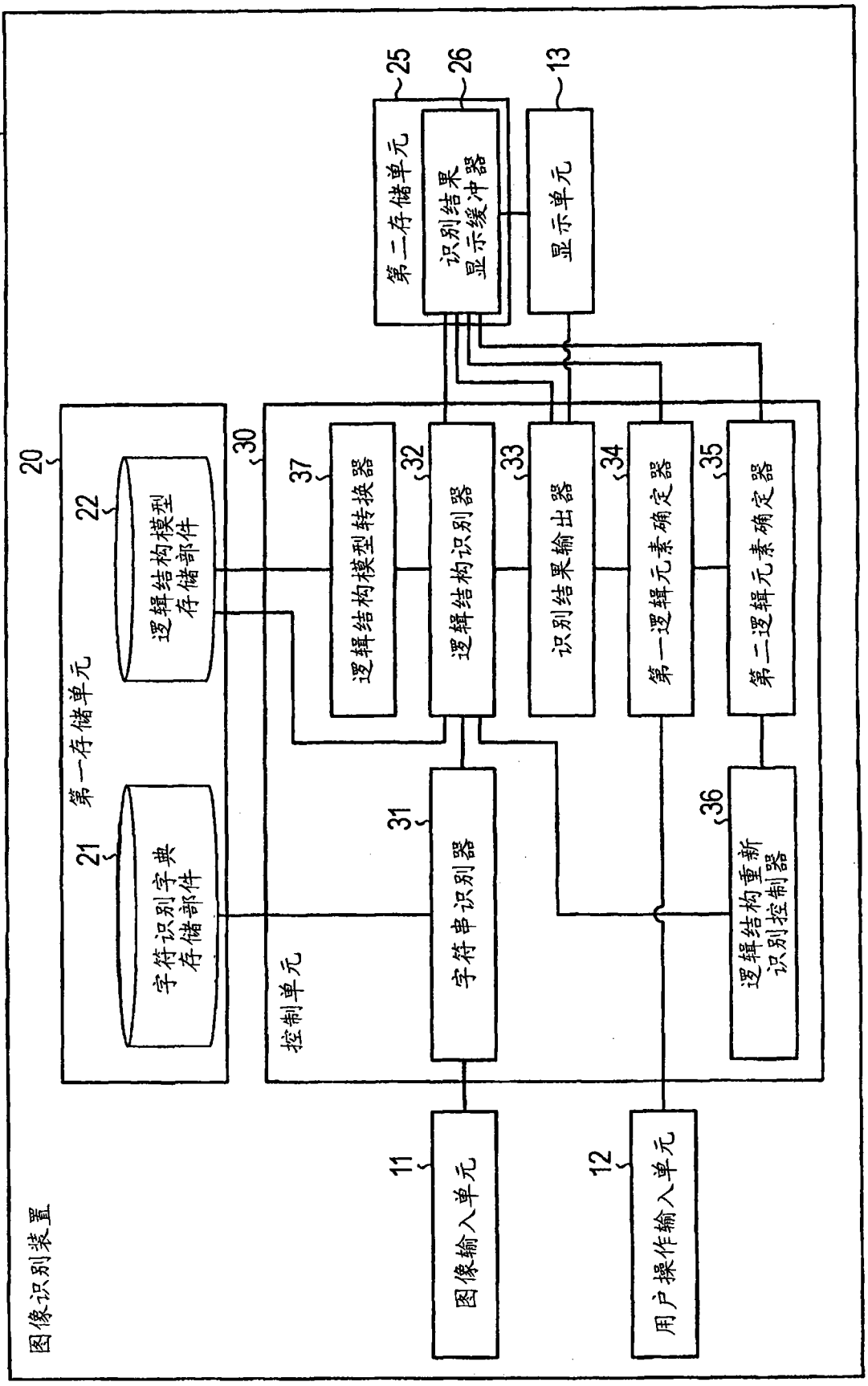

10
图像识别装置
11
图像输入单元
12
用户操作输入单元
13
显示单元
20
第一存储单元
21
字符识别字典存储部件
22
逻辑结构模型存储部件
25
第二存储单元
26
识别结果显示缓冲器
30
控制单元
31
字符串识别器
32
逻辑结构识别器
33
识别结果输出器
34
第一逻辑元素确定器
35
第二逻辑元素确定器
36
逻辑结构重新识别控制器
37
逻辑结构模型转换器

图 2

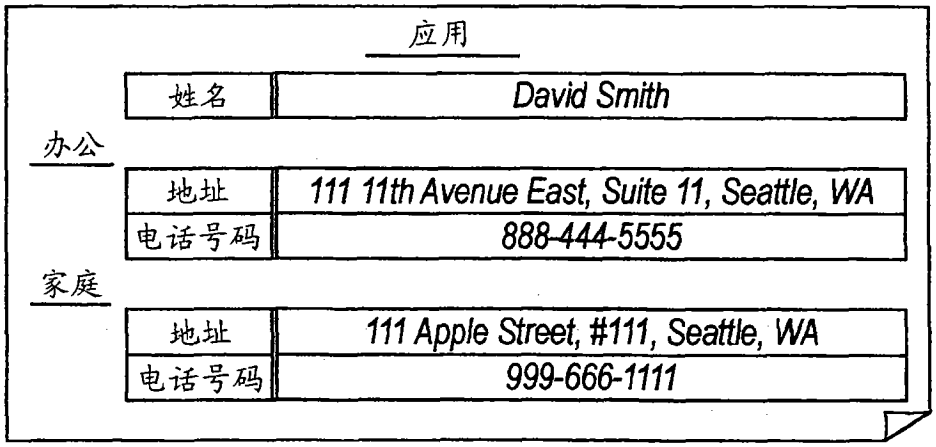

应用
姓名
David Smith
办公
地址
111 11th Avenue East, Suite 11, Seattle, WA
电话号码
888-444-5555
家庭
地址
111 Apple Street, #111, Seattle, WA
电话号码
999-666-1111

图 3

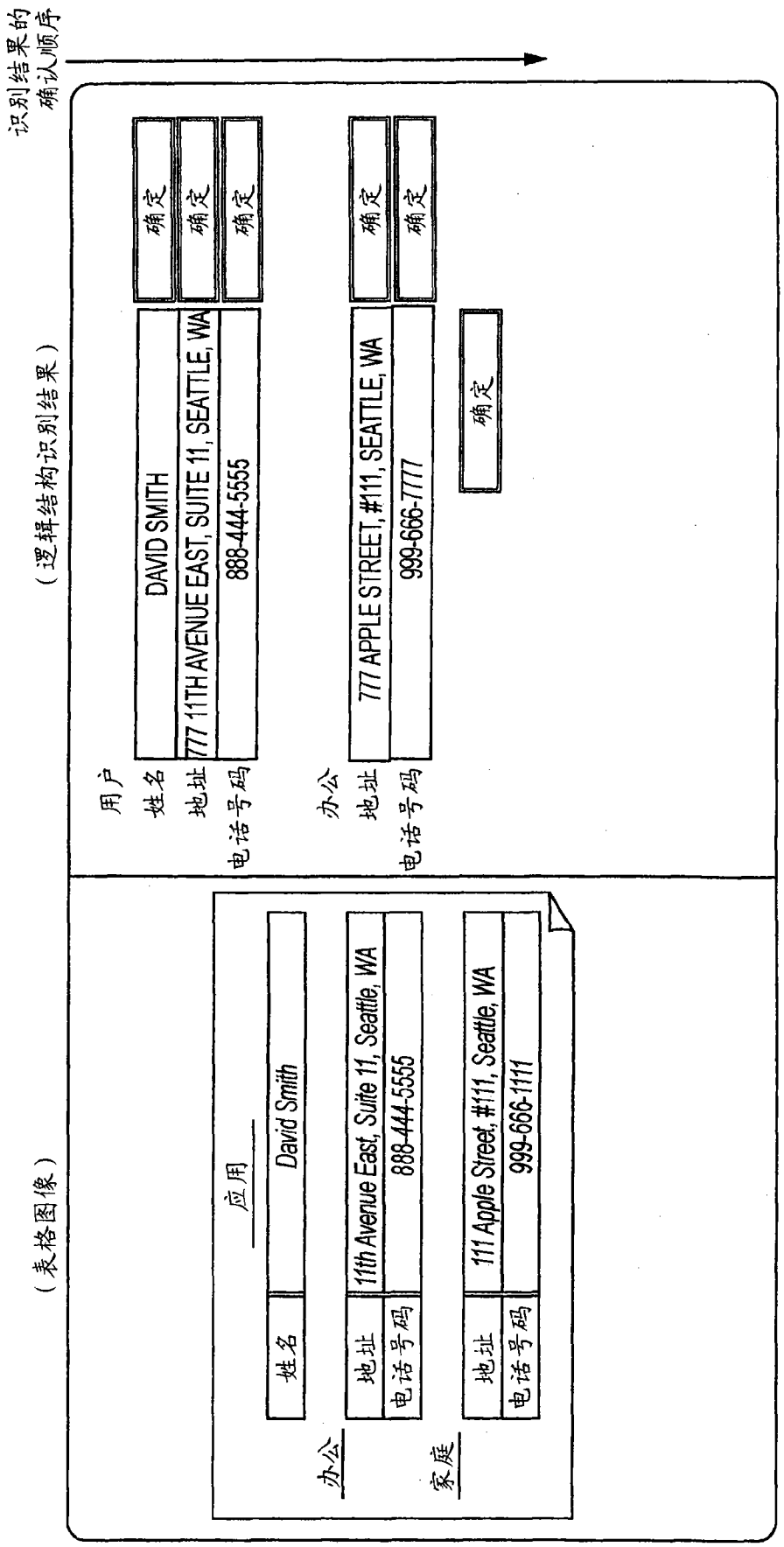

识别结果的
确认顺序
（逻辑结构识别结果）
用户
姓名
地址
电话号码
DAVID SMITH
777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
888-444-5555
确定
确定
确定
办公
地址
电话号码
777 APPLE STREET, #111, SEATTLE, WA
999-666-7777
确定
确定
确定
（表格图像）
应用
姓名
David Smith
办公
地址
11th Avenue East, Suite 11, Seattle, WA
电话号码
888-444-5555
家庭
地址
111 Apple Street, #111, Seattle, WA
电话号码
999-666-1111

图 4

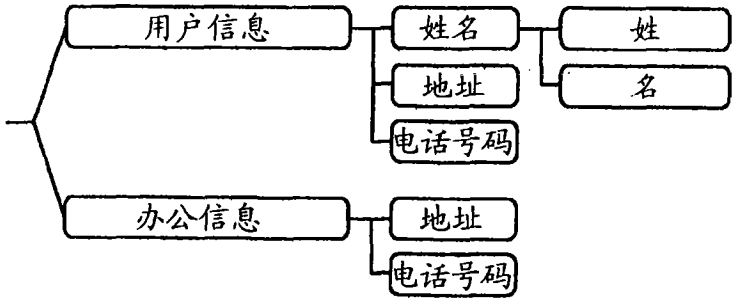


图 5A

ID		层次	父节点	位置	属性	数据		
1	用户信息	0	0	7:下	组			
2	姓名	1	1		组			
3	地址	1	1	4:右，下	节点	字符串	数字串	"-"
4	电话号码	1	1	3:左，上	节点	数字串	"-"	
5	姓	2	2	6:右	节点	字符串		
6	名	2	2	5:左	节点	字符串		
7	办公信息	0	0	1:上	组			
8	地址	1	7	9:右，下	节点	字符串	数字串	"-"
9	电话号码	1	7	8:左，上	节点	数字串	"-"	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

图 5B

逻辑元素的类型			针对字符串的候选			
			第一候选	第二候选	第三候选	第四候选
用户信息	姓名	姓	SMITH	-	-	-
		名	DAVID	DEVID	-	-
	地址		777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	777 APPLE STREET, #111, SEATTLE, WA	111 APPLE STREET, #111, SEATTLE, WA
	电话号码		888-444-5555	888-111-5555	999-666-7777	999-666-1111
办公信息	地址		777 APPLE STREET, #111, SEATTLE, WA	111 APPLE STREET, #111, SEATTLE, WA	777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
	电话号码		999-666-7777	999-666-1111	888-444-5555	888-111-5555

图 6

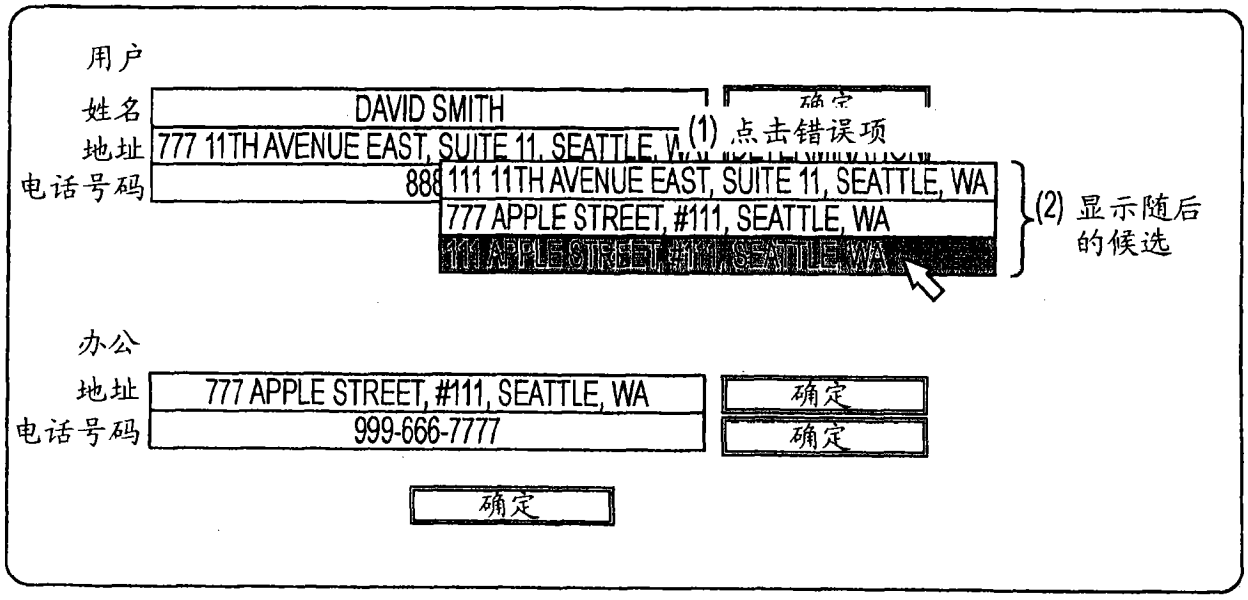

用户
姓名
DAVID SMITH
确定
地址
777 11TH AVENUE EAST, SUITE 11, SEATTLE, W
(1) 点击错误项
电话号码
888
111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
777 APPLE STREET, #111, SEATTLE, WA
111 APPLE STREET #111 SEATTLE WA
(2) 显示随后的候选
办公
地址
777 APPLE STREET, #111, SEATTLE, WA
确定
电话号码
999-666-7777
确定
确定

图 7

逻辑元素的类型			针对字符串的候选			
			第一候选	第二候选	第三候选	第四候选
用户信息	姓名	姓	SMITH	-	-	-
		名	DAVID	-	-	-
	地址		111 APPLE STREET #111, SEATTLE, WA	-	-	-
	电话号码		888-444-5555	888-111-5555	999-666-7777	999-666-1111
办公信息	地址		777 APPLE STREET, #111, SEATTLE, WA	111 APPLE STREET, #111, SEATTLE, WA	777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
	电话号码		999-666-7777	999-666-1111	888-444-5555	888-111-5555

图 8A

逻辑元素的类型			针对字符串的候选			
			第一候选	第二候选	第三候选	第四候选
用户信息	姓名	姓	SMITH	-	-	-
		名	DAVID	-	-	-
	地址		111 APPLE STREET #111, SEATTLE, WA	-	-	-
	电话号码		999-666-7777	999-666-1111	-	-
办公信息	地址		777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	-	-
	电话号码		888-444-5555	888-111-5555	-	-

图 8B

逻辑元素的类型			针对字符串的候选			
			第一候选	第二候选	第三候选	第四候选
用户信息	姓名	姓	SMITH	-	-	-
		名	DAVID	-	-	-
	地址		1 APPLE STREET #11, SEATTLE, WA	-	-	-
	电话号码		999-666-1111	-	-	-
办公信息	地址		777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	-	-
	电话号码		888-444-5555	888-111-5555	-	-

图 8C

逻辑元素的类型			针对字符串的候选			
			第一候选	第二候选	第三候选	第四候选
用户信息	姓名	姓	SMITH	-	-	-
		名	DAVID	-	-	-
	地址		1 APPLE STREET #11, SEATTLE, WA	-	-	-
	电话号码		999-666-1111		-	-
办公信息	地址		777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	-	-
	电话号码		888-444-5555	888-111-5555	-	-

图 8D

ID		层次	父节点	位置	属性	数据		
1	用户信息	0	0	7:下	组			
2	姓名	1	1		组			
3	地址	1	1	4:右，下	节点	111 APPLE STREET, #111, SEATTLE, WA		
4	电话号码	1	1	3:左，上	节点	数字串	"-"	
5	姓	2	2	6:右	节点	SMITH		
6	名	2	2	5:左	节点	DAVID		
7	办公信息	0	0	1:上	组			
8	地址	1	7	9:右，下	节点	字符串	数字串	"-"
9	电话号码	1	7	8:左，上	节点	数字串	"-"	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

图 9A

ID		层次	父节点	位置	属性	数据		
1	用户信息	0	0	7:下	组			
2	姓名	1	1		组			
3	地址	1	1	4:右，下	节点	111 APPLE STREET, #111, SEATTLE, WA		
4	电话号码	1	1	3:左，上	节点	999-666-1111		
5	姓	2	2	6:右	节点	SMITH		
6	名	2	2	5:左	节点	DAVID		
7	办公信息	0	0	1:上	组			
8	地址	1	7	9:右，下	节点	字符串	数字串	"-"
9	电话号码	1	7	8:左，上	节点	数字串	"-"	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

图 9B

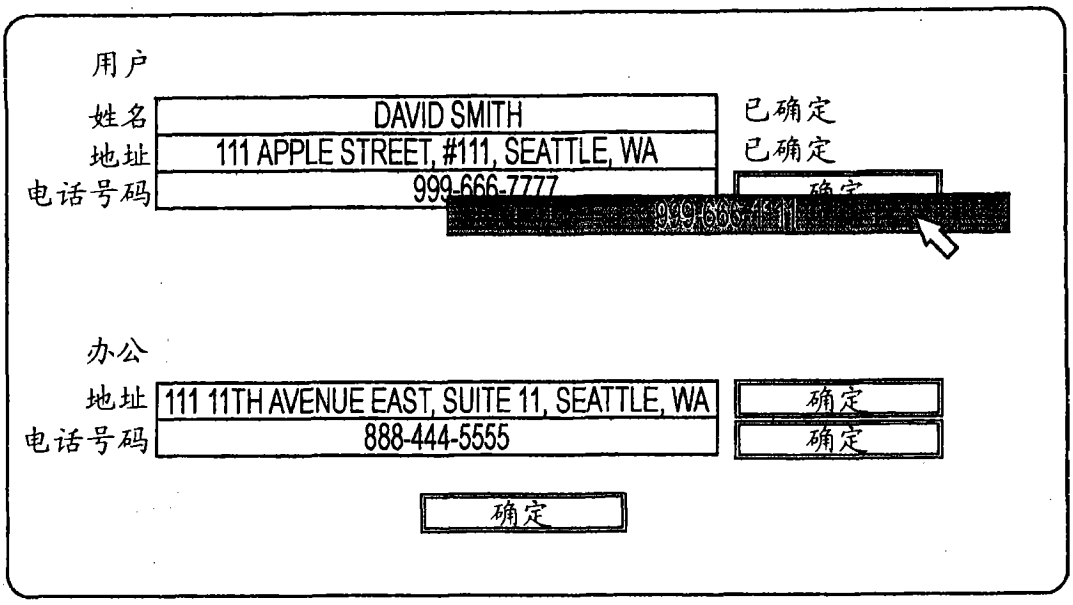

用户
姓名
DAVID SMITH
已确定
地址
111 APPLE STREET, #111, SEATTLE, WA
已确定
电话号码
999-666-7777
确定
办公
地址
111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
确定
电话号码
888-444-5555
确定
确定

图 10A

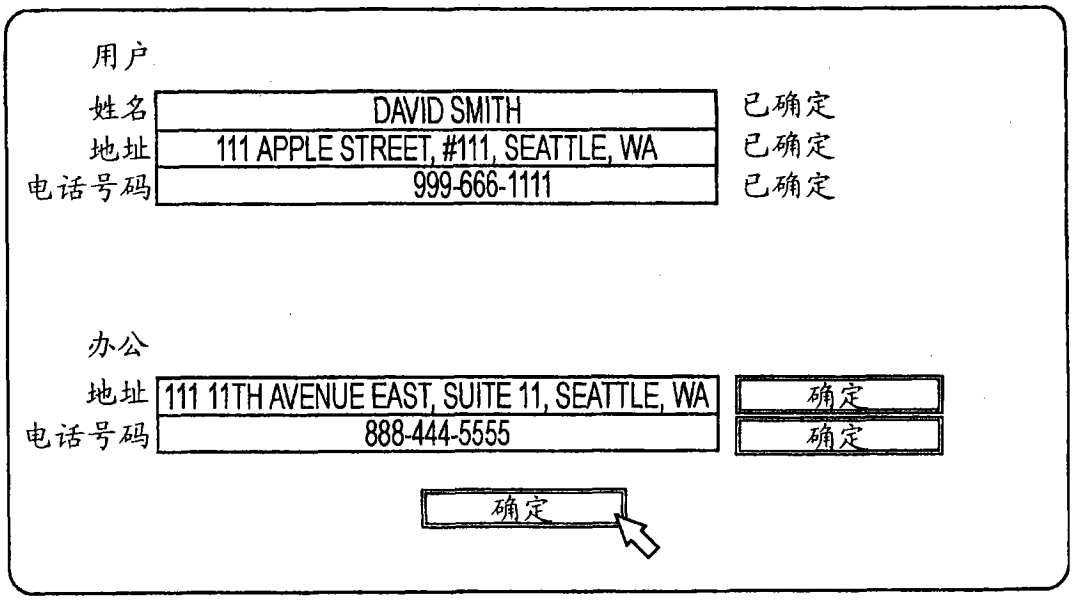

用户
姓名
DAVID SMITH
已确定
地址
111 APPLE STREET, #111, SEATTLE, WA
已确定
电话号码
999-666-1111
已确定
办公
地址
111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
确定
电话号码
888-444-5555
确定
确定

图 10B

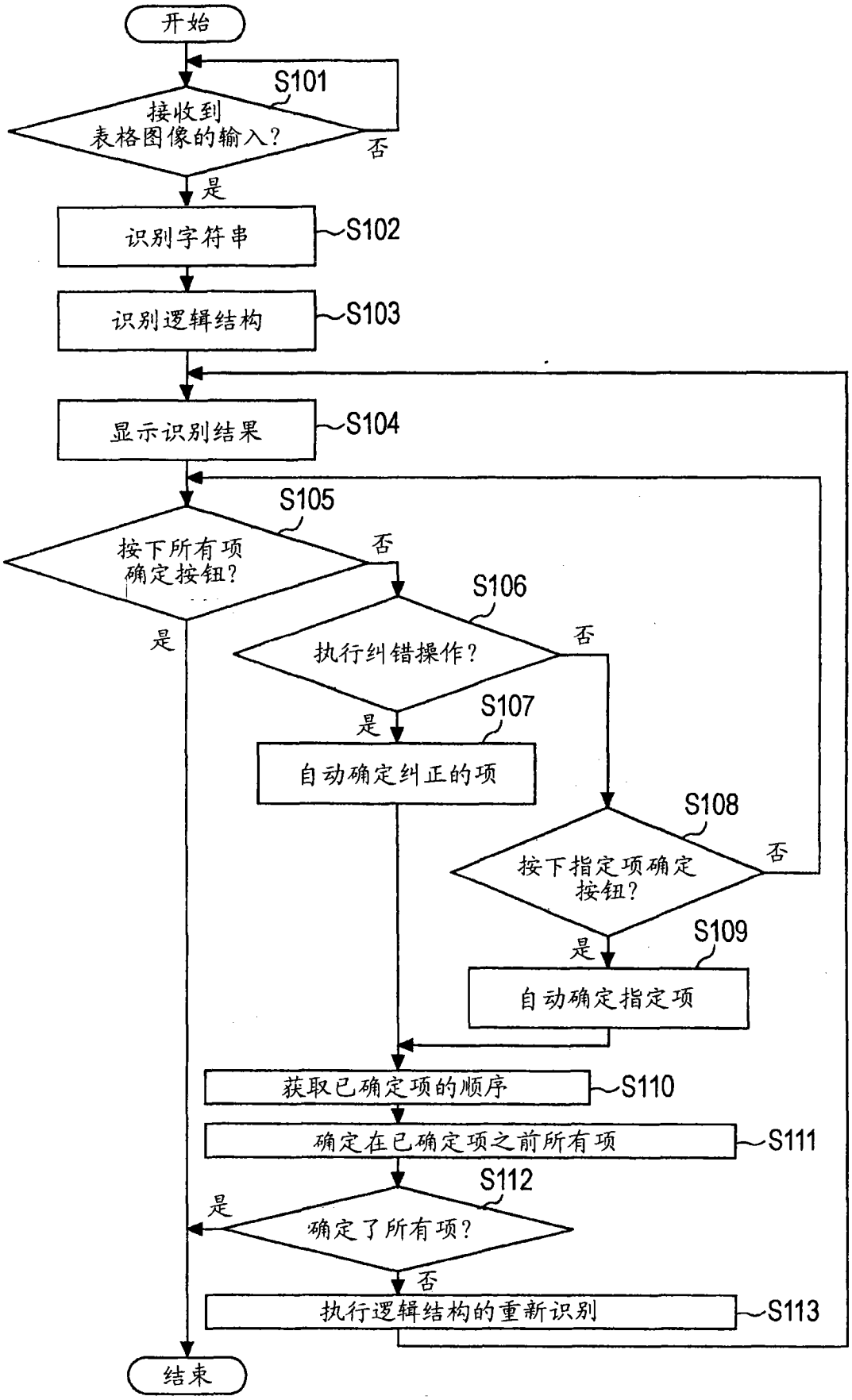

开始
S101
接收到表格图像的输入?
否
是
识别字符串
S102
识别逻辑结构
S103
显示识别结果
S104
S105
按下所有项确定按钮?
否
是
S106
执行纠错操作?
否
是
S107
自动确定纠正的项
S108
按下指定项确定按钮?
否
是
S109
自动确定指定项
获取已确定项的顺序
S110
确定在已确定项之前所有项
S111
S112
确定了所有项?
是
否
执行逻辑结构的重新识别
S113
结束

图 11

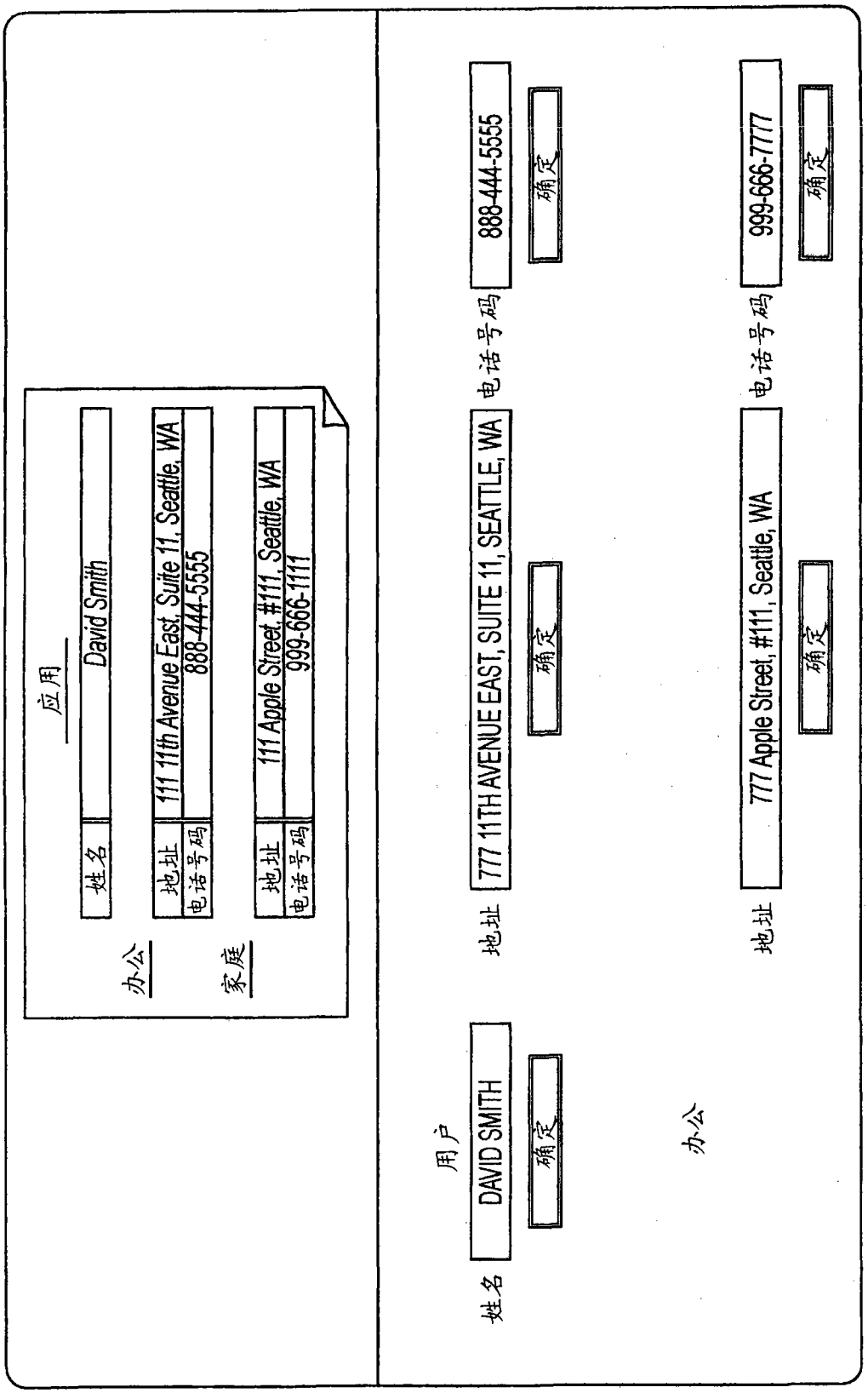

应用
姓名
David Smith
办公
地址
111 11th Avenue East, Suite 11, Seattle, WA
电话号码
888-444-5555
家庭
地址
111 Apple Street #111, Seattle, WA
电话号码
999-666-1111
用户
姓名
DAVID SMITH
确定
地址
777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
确定
电话号码
888-444-5555
确定
办公
地址
777 Apple Street, #111, Seattle, WA
确定
电话号码
999-666-7777
确定

图 12

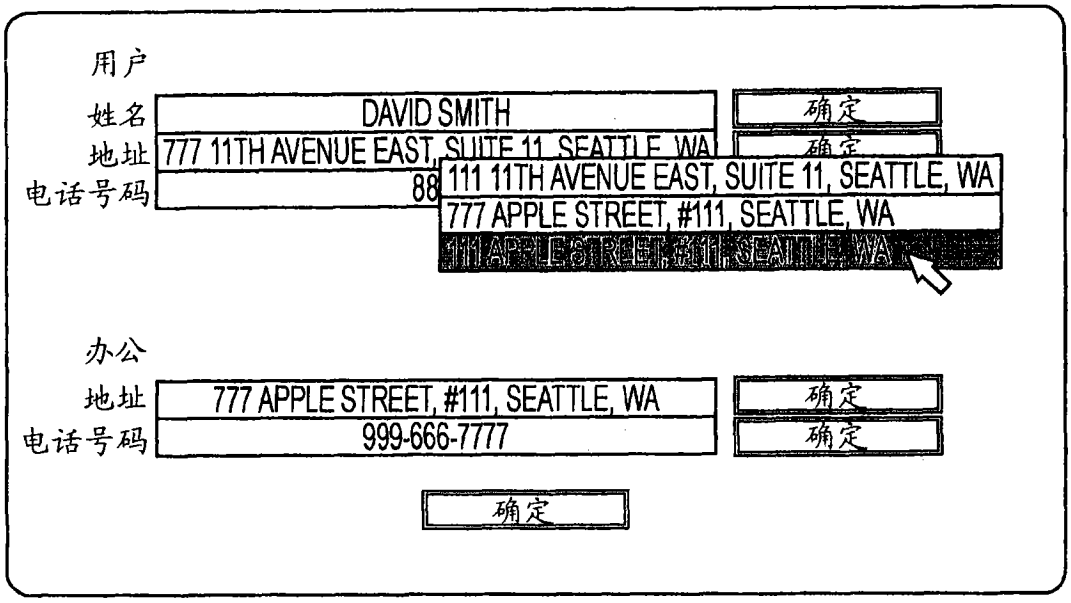

用户
姓名
DAVID SMITH
确定
地址
777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
确定
电话号码
888
111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
777 APPLE STREET, #111, SEATTLE, WA
111 APPLE STREET, #111, SEATTLE, WA
办公
地址
777 APPLE STREET, #111, SEATTLE, WA
确定
电话号码
999-666-7777
确定
确定

图 13A

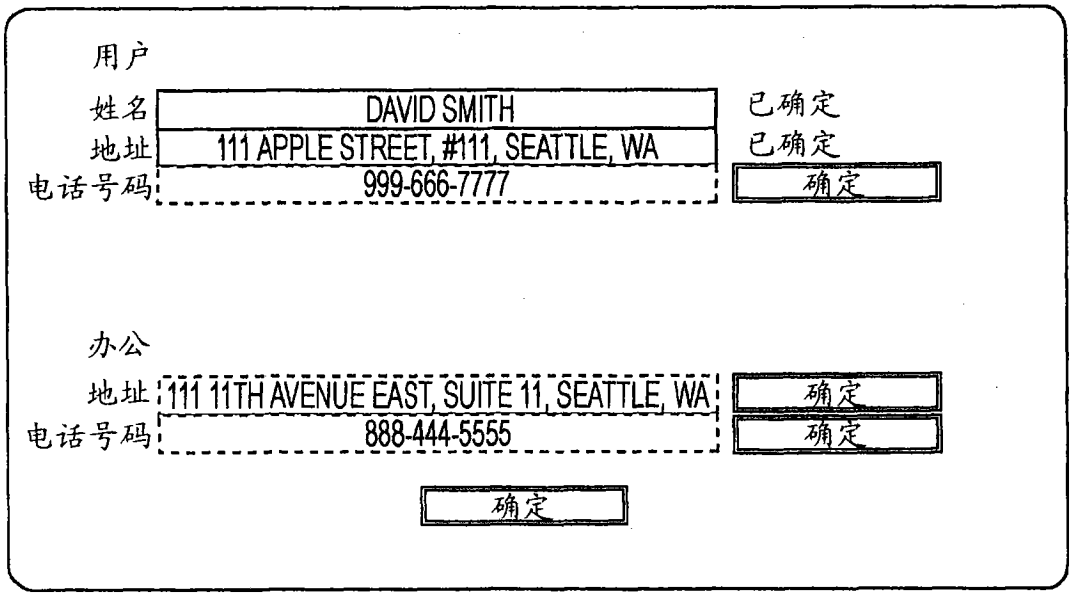

用户
姓名
DAVID SMITH
已确定
地址
111 APPLE STREET, #111, SEATTLE, WA
已确定
电话号码
999-666-7777
确定
办公
地址
111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
确定
电话号码
888-444-5555
确定
确定

图 13B

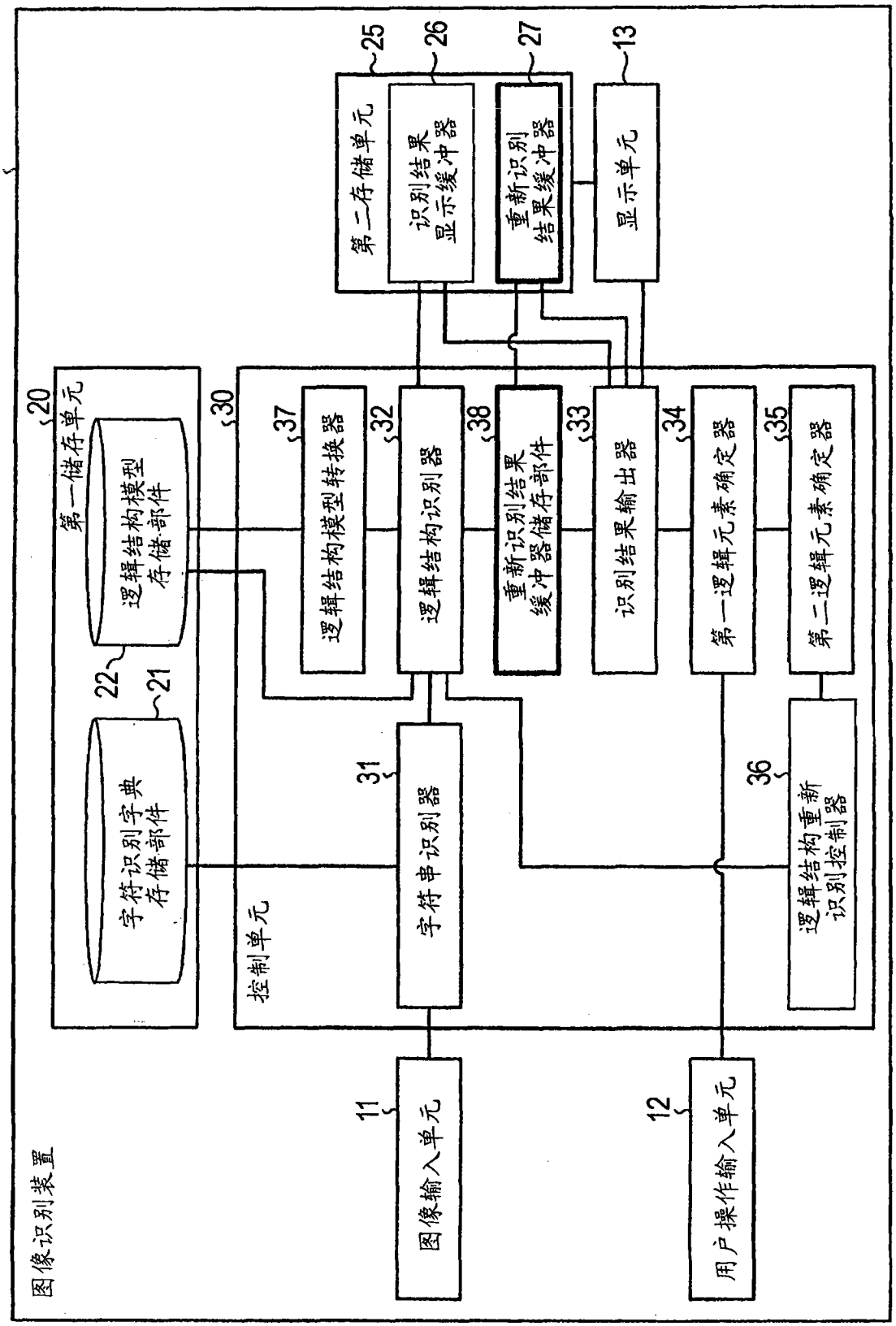

10
图像识别装置
第一储存单元
20
逻辑结构模型存储部件
22
字符识别字典存储部件
21
控制单元
30
逻辑结构模型转换器
37
逻辑结构识别器
32
重新识别结果缓冲器储存部件
38
识别结果输出器
33
第一逻辑元素确定器
34
第二逻辑元素确定器
35
字符串识别器
31
逻辑结构重新识别控制器
36
图像输入单元
11
用户操作输入单元
12
第二存储单元
25
识别结果显示缓冲器
26
重新识别结果缓冲器
27
显示单元
13

图 14

逻辑元素的类型			针对字符串的候选			
			第一候选	第二候选	第三候选	第四候选
用户信息	姓名	姓	SMITH	-	-	-
		名	DAVID	-	-	-
	地址		777 APPLE STREET, #111, SEATTLE, WA	-	-	-
	电话号码		888-444-5555	888-111-5555	999-666-7777	999-666-1111
办公信息	地址		777 APPLE STREET, #111, SEATTLE, WA	111 APPLE STREET, #111, SEATTLE, WA	777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
	电话号码		999-666-7777	999-666-1111	888-444-5555	888-111-5555

图 15A

逻辑元素的类型			针对字符串的候选			
			第一候选	第二候选	第三候选	第四候选
用户信息	姓名	姓	SMITH	-	-	-
		名	DAVID	-	-	-
	地址		777 APPLE STREET, #111, SEATTLE, WA	-	-	-
	电话号码		999-666-7777 (HIGHLIGHTED)	999-666-1111	-	-
办公信息	地址		111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA (HIGHLIGHTED)	777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	-	-
	电话号码		888-444-5555 (HIGHLIGHTED)	888-111-5555	-	-

图 15B

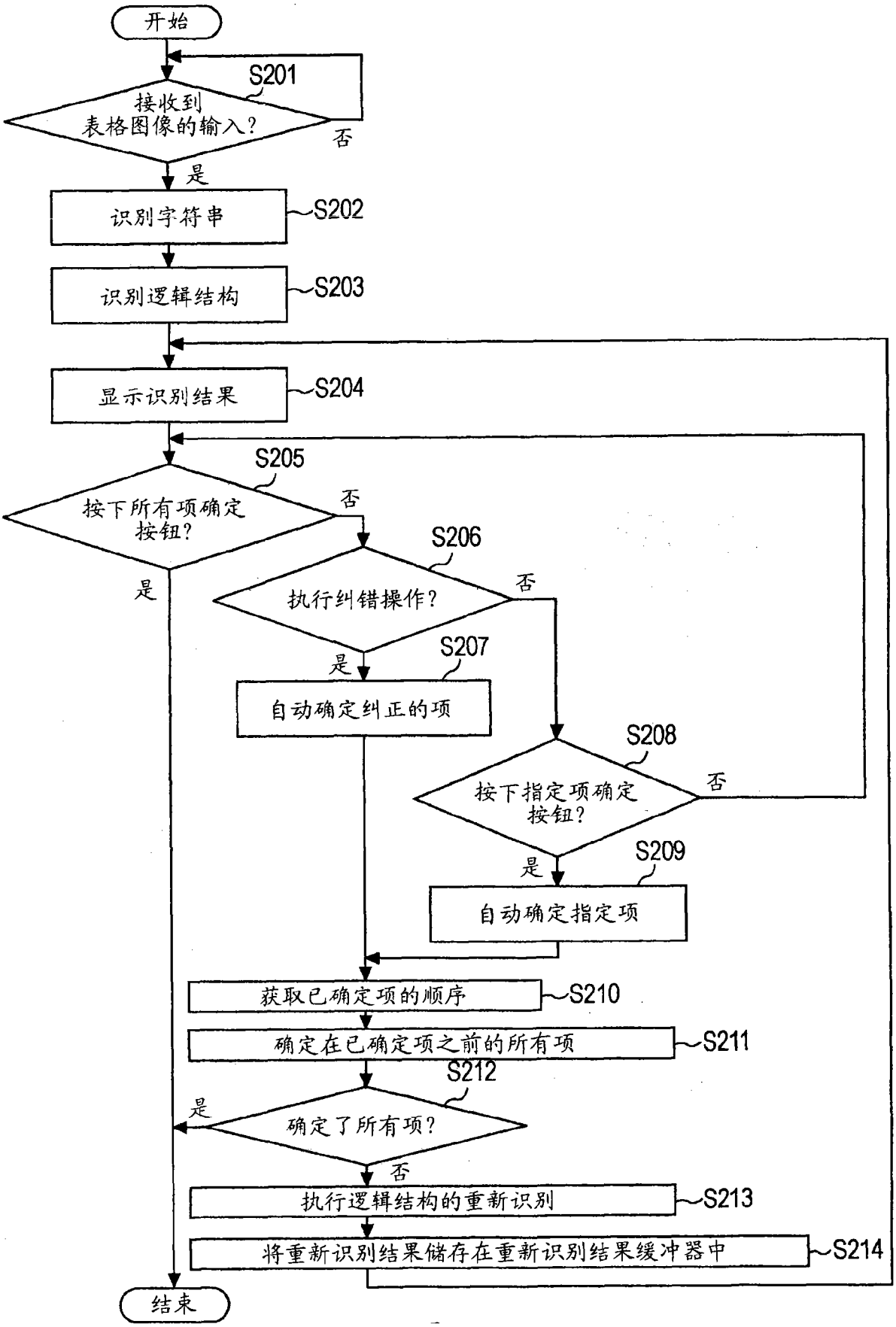

开始
S201
接收到表格图像的输入?
否
是
识别字符串
S202
识别逻辑结构
S203
显示识别结果
S204
S205
按下所有项确定按钮?
否
是
S206
执行纠错操作?
否
是
S207
自动确定纠正的项
S208
按下指定项确定按钮?
否
是
S209
自动确定指定项
获取已确定项的顺序
S210
确定在已确定项之前的所有项
S211
S212
确定了所有项?
是
否
执行逻辑结构的重新识别
S213
将重新识别结果储存在重新识别结果缓冲器中
S214
结束

图 16

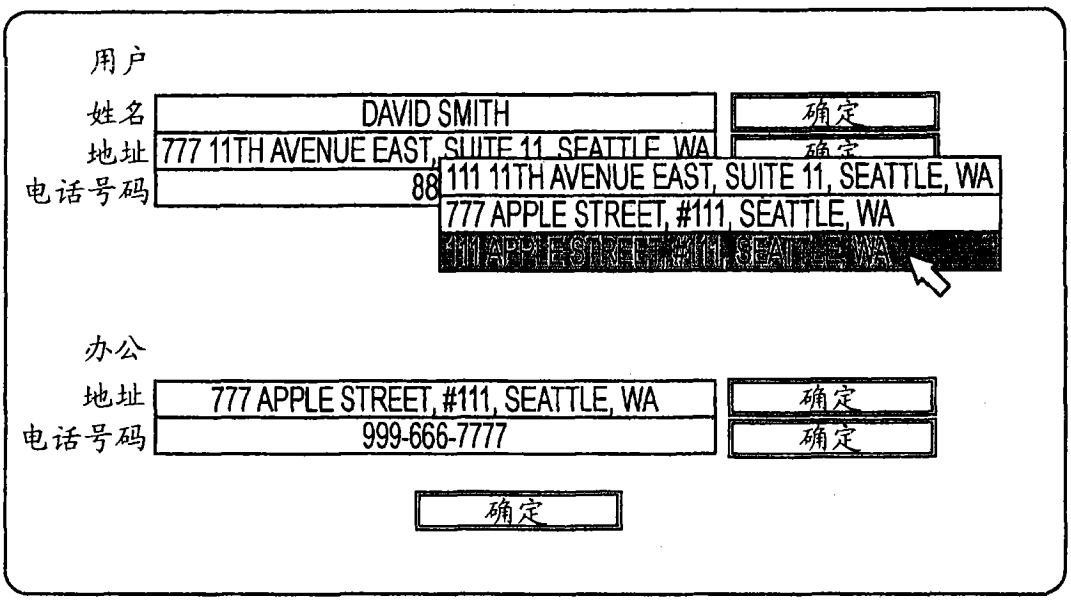

用户
姓名
DAVID SMITH
确定
地址
777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
确定
电话号码
888
111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
777 APPLE STREET, #111, SEATTLE, WA
111 APPLE STREET, #111, SEATTLE, WA
办公
地址
777 APPLE STREET, #111, SEATTLE, WA
确定
电话号码
999-666-7777
确定
确定

图 17A

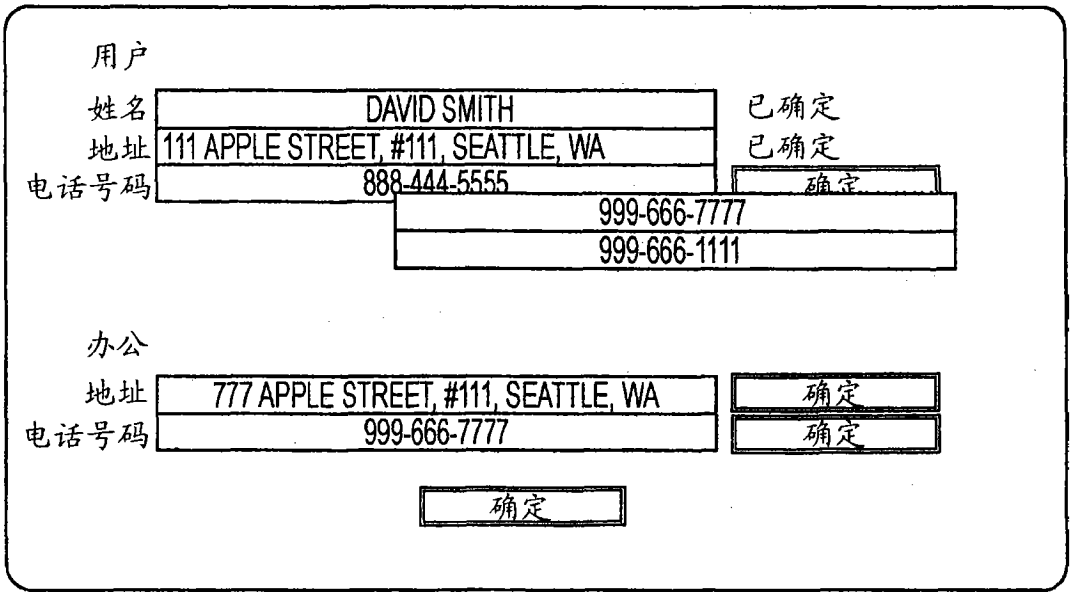

用户
姓名
DAVID SMITH
已确定
地址
111 APPLE STREET, #111, SEATTLE, WA
已确定
电话号码
888-444-5555
确定
999-666-7777
999-666-1111
办公
地址
777 APPLE STREET, #111, SEATTLE, WA
确定
电话号码
999-666-7777
确定
确定

图 17B

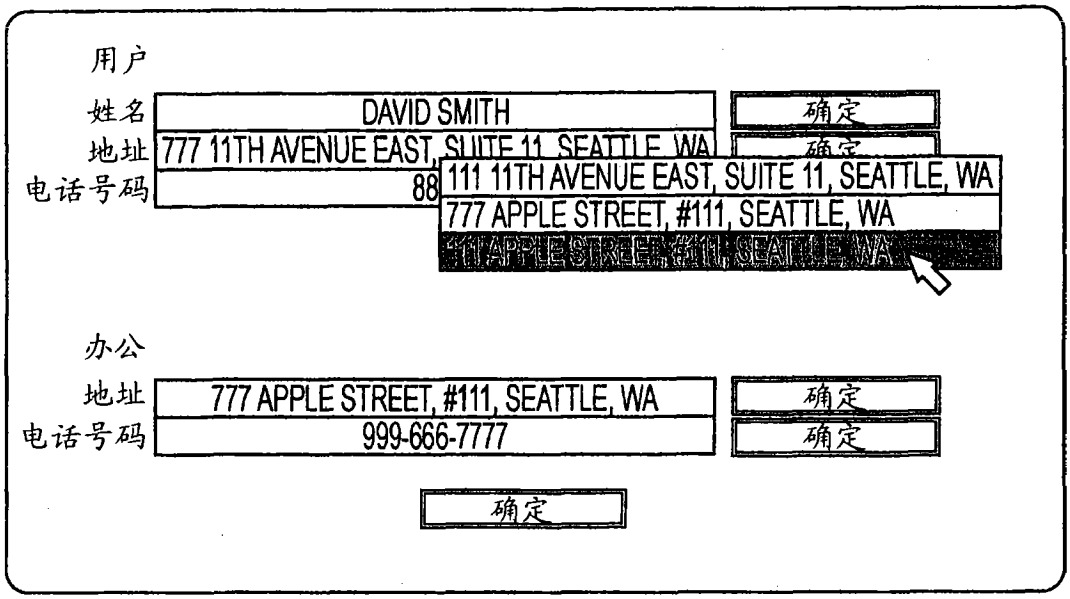

用户
姓名
DAVID SMITH
确定
地址
777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
确定
电话号码
888
111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
777 APPLE STREET, #111, SEATTLE, WA
办公
地址
777 APPLE STREET, #111, SEATTLE, WA
确定
电话号码
999-666-7777
确定
确定

图 18A

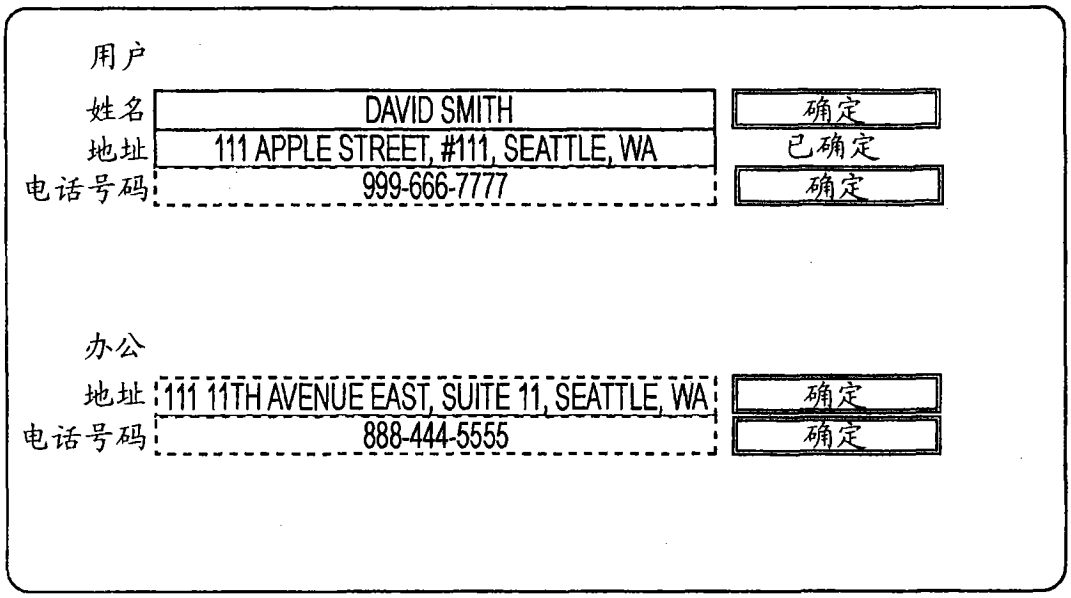

用户
姓名
DAVID SMITH
确定
地址
111 APPLE STREET, #111, SEATTLE, WA
已确定
电话号码
999-666-7777
确定
办公
地址
111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
确定
电话号码
888-444-5555
确定

图 18B

逻辑元素的类型			针对字符串的候选			
			第一候选	第二候选	第三候选	第四候选
用户信息	姓名	姓	SMITH	-	-	-
		第一个名	DAVID	DEVID	-	-
	地址		111 APPLE STREET, #111, SEATTLE, WA	-	-	-
	电话号码		888-444-5555	888-111-5555	999-666-7777	999-666-1111
办公信息	地址		777 APPLE STREET, #111, SEATTLE, WA	111 APPLE STREET, #111, SEATTLE, WA	777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
	电话号码		999-666-7777	999-666-1111	888-444-5555	888-111-5555

图 19A

逻辑元素的类型			针对字符串的候选			
			第一候选	第二候选	第三候选	第四候选
用户信息	姓名	姓	SMITH	-	-	-
		第一个名	DAVID	DEVID	-	-
	地址		111 APPLE STREET, #111, SEATTLE, WA	-	-	-
	电话号码		999-666-7777 (HIGHLIGHTED)	999-666-1111	-	-
办公信息	地址		111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA (HIGHLIGHTED)	777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA	-	-
	电话号码		888-444-5555 (HIGHLIGHTED)	888-111-5555	-	-

图 19B

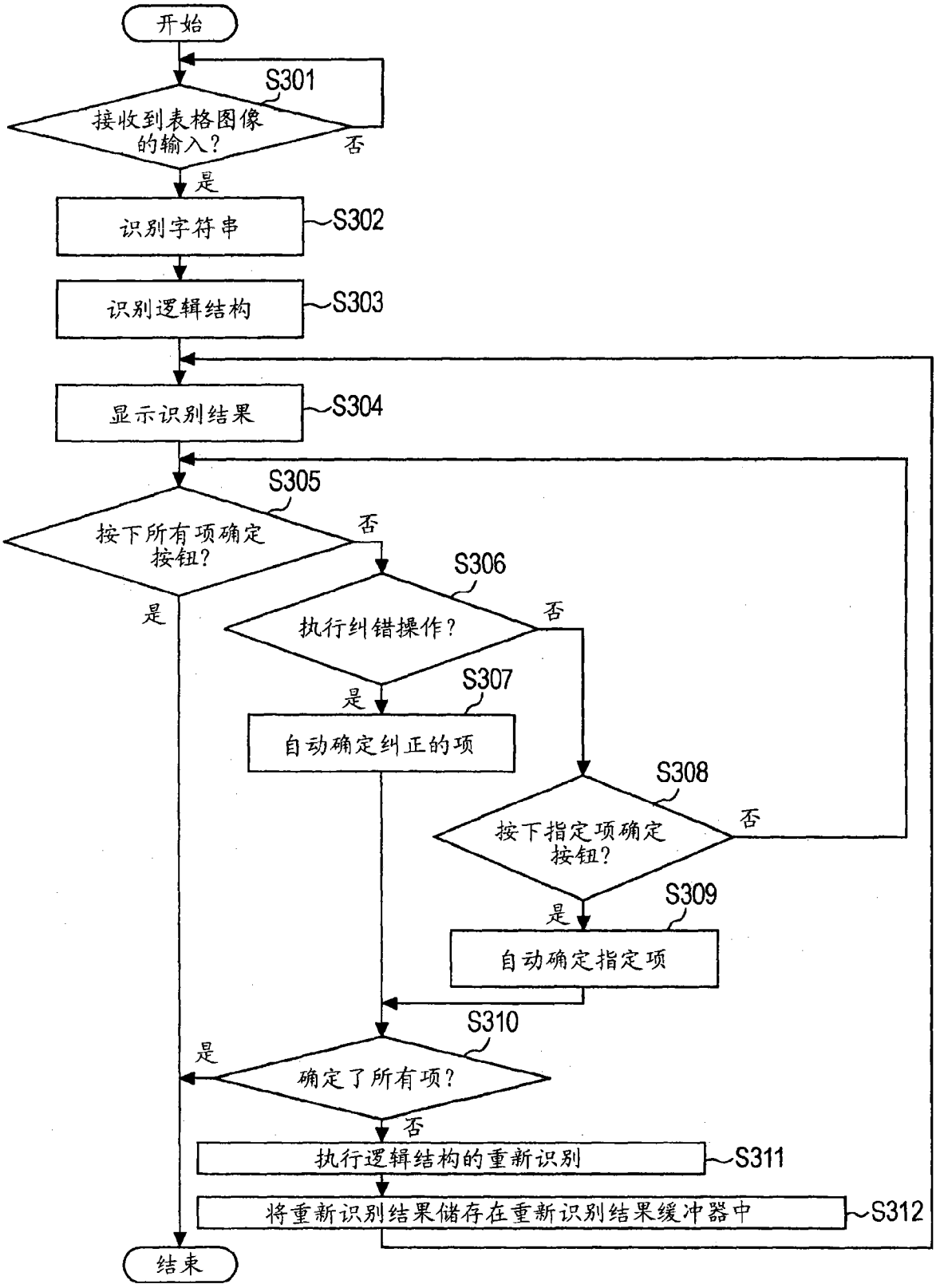

开始
S301
接收到表格图像的输入?
否
是
识别字符串
S302
识别逻辑结构
S303
显示识别结果
S304
S305
按下所有项确定按钮?
否
是
S306
执行纠错操作?
否
是
S307
自动确定纠正的项
S308
按下指定项确定按钮?
否
是
S309
自动确定指定项
S310
确定了所有项?
是
否
执行逻辑结构的重新识别
S311
将重新识别结果储存在重新识别结果缓冲器中
S312
结束

图 20

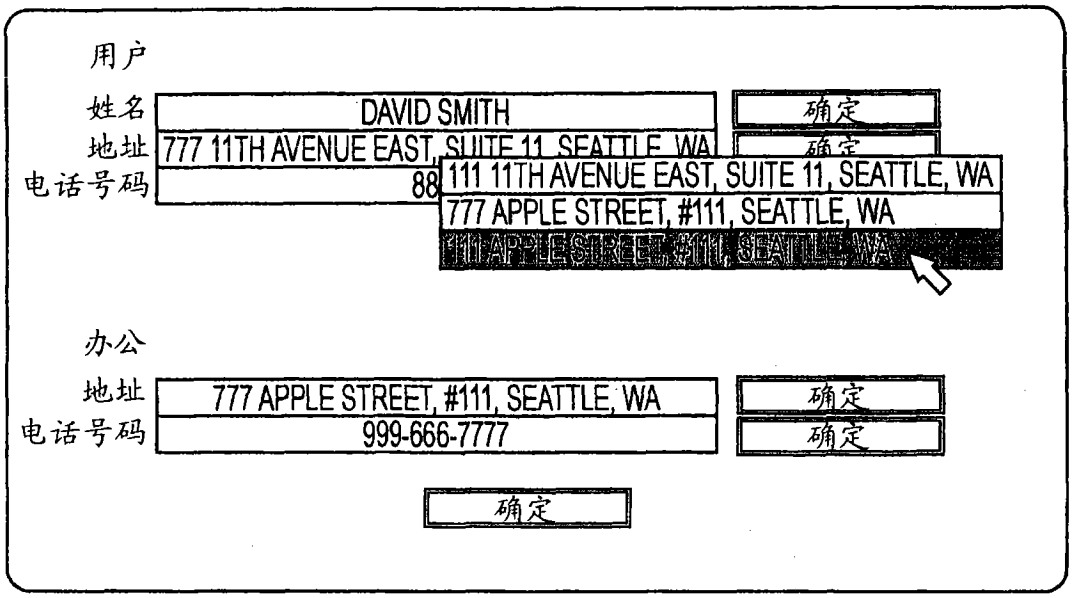

用户
姓名
DAVID SMITH
确定
地址
777 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
确定
电话号码
888
111 11TH AVENUE EAST, SUITE 11, SEATTLE, WA
777 APPLE STREET, #111, SEATTLE, WA
111 APPLE STREET, #111, SEATTLE, WA
办公
地址
777 APPLE STREET, #111, SEATTLE, WA
确定
电话号码
999-666-7777
确定
确定

图 21A

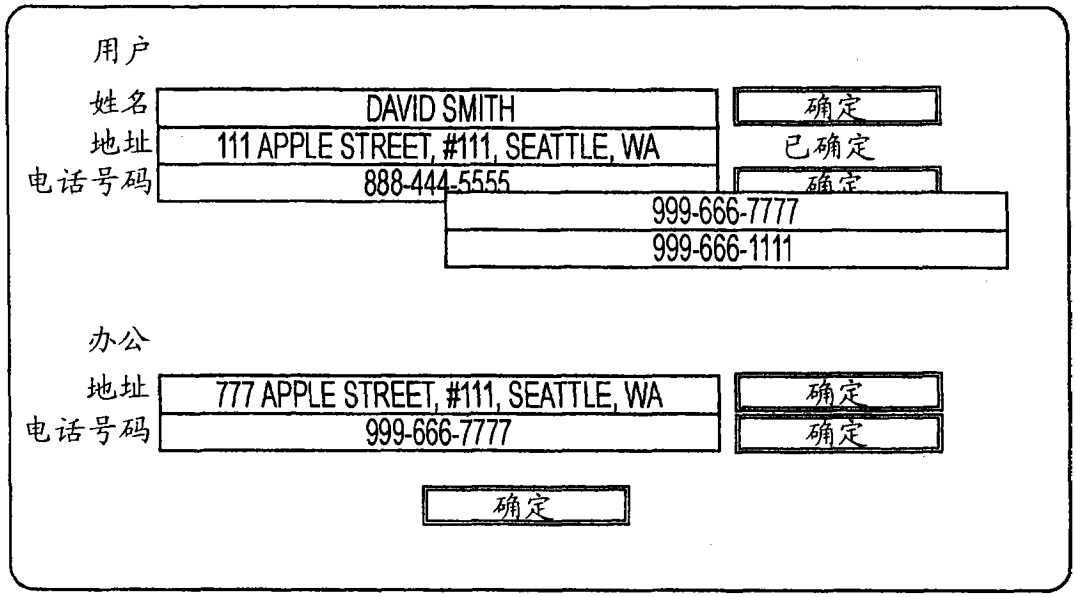

用户
姓名
DAVID SMITH
确定
地址
111 APPLE STREET, #111, SEATTLE, WA
已确定
电话号码
888-444-5555
确定
999-666-7777
999-666-1111
办公
地址
777 APPLE STREET, #111, SEATTLE, WA
确定
电话号码
999-666-7777
确定
确定

图 21B

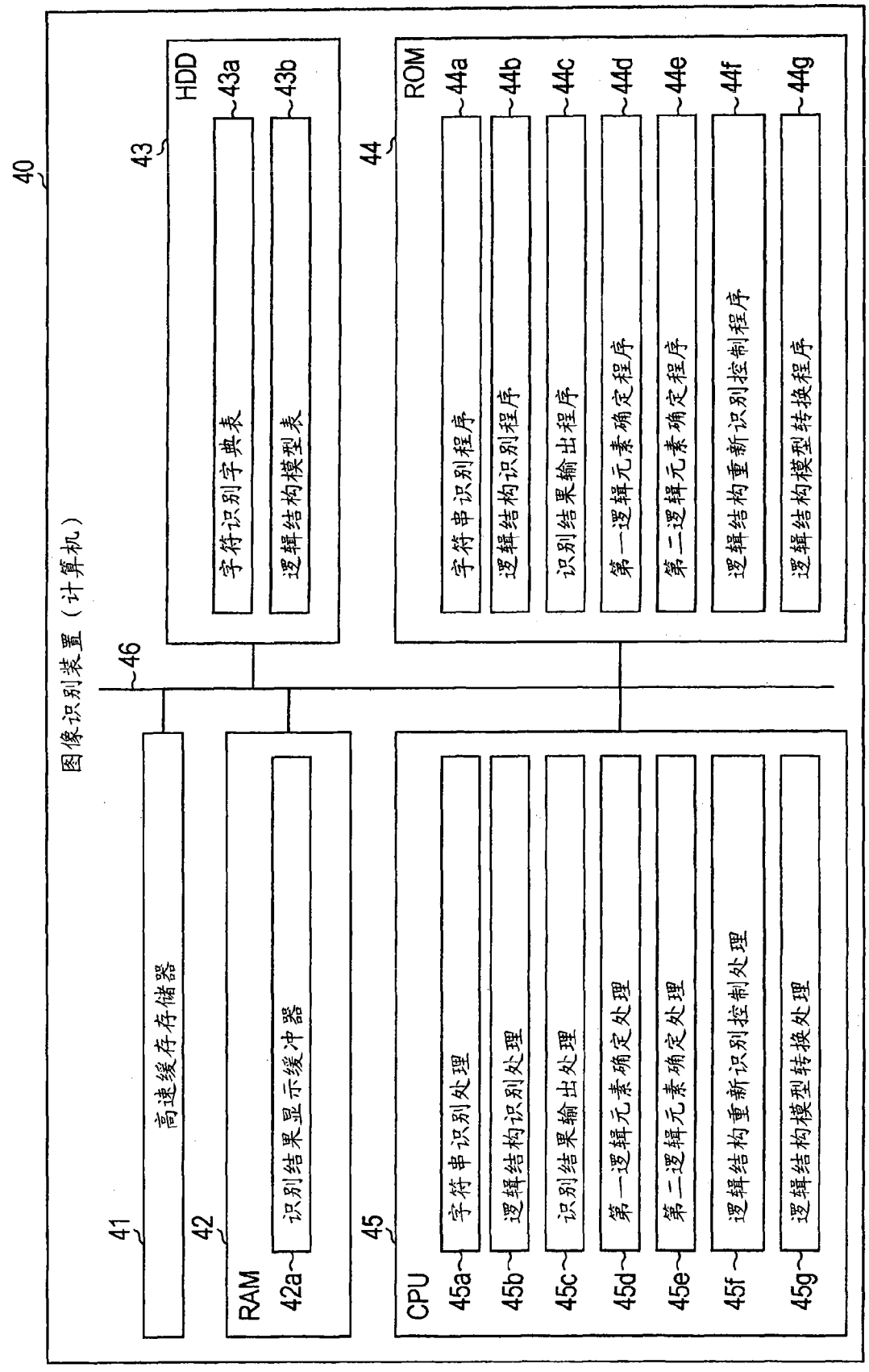

图像识别装置（计算机）
40
HDD
43
43a
字符识别字典表
43b
逻辑结构模型表
ROM
44
44a
字符串识别程序
44b
逻辑结构识别程序
44c
识别结果输出程序
44d
第一逻辑元素确定程序
44e
第二逻辑元素确定程序
44f
逻辑结构重新识别控制程序
44g
逻辑结构模型转换程序
46
高速缓存存储器
41
RAM
42
42a
识别结果显示缓冲器
CPU
45
45a
字符串识别处理
45b
逻辑结构识别处理
45c
识别结果输出处理
45d
第一逻辑元素确定处理
45e
第二逻辑元素确定处理
45f
逻辑结构重新识别控制处理
45g
逻辑结构模型转换处理

图 22