

# 發明專利說明書

200416583

(本申請書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※申請案號：92132932

※申請日期：92年11月24日

※IPC分類：

G06F19/00

## 壹、發明名稱：

(中) 帳簿傳票之定義資料作成方法及帳簿傳票之處理裝置

(外) 帳票定義データ作成方法および帳票処理装置

## 貳、申請人：(共1人)

1. 姓名：(中) 日立製作所股份有限公司

(英) HITACHI, LTD.

代表人：(中) 1. 庄山悦彦

(英)

地址：(中) 日本國東京都千代田區神田駿河台四丁目六番地

(英)

國籍：(中英) 日本

JAPAN

## 參、發明人：(共2人)

1. 姓名：(中) 淺野英輔

(英) 浅野英輔

地址：(中) 日本國東京都千代田區丸之内一丁目五番一號新丸大樓日立製作所  
(股) 知的所有權本部內

(英) 日本国東京都千代田区丸の内一丁目5番1号新丸ビル(株)日立製作所 知的所有權本部內

2. 姓名：(中) 新庄廣

(英) 新庄広

地址：(中) 日本國東京都千代田區丸之内一丁目五番一號新丸大樓日立製作所  
(股) 知的所有權本部內

(英) 日本国東京都千代田区丸の内一丁目5番1号新丸ビル(株)日立製作所 知的所有權本部內

## 肆、聲明事項：

◎本案申請前已向下列國家(地區)申請專利  主張國際優先權：

【格式請依：受理國家(地區)；申請日；申請案號數 順序註記】

# 發明專利說明書

200416583

(本申請書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※申請案號：92132932

※申請日期：92年11月24日

※IPC分類：

G06F19/00

## 壹、發明名稱：

(中) 帳簿傳票之定義資料作成方法及帳簿傳票之處理裝置

(外) 帳票定義データ作成方法および帳票処理装置

## 貳、申請人：(共 1 人)

1. 姓名：(中) 日立製作所股份有限公司

(英) HITACHI, LTD.

代表人：(中) 1. 庄山悦彦

(英)

地址：(中) 日本國東京都千代田區神田駿河台四丁目六番地

(英)

國籍：(中英) 日本

JAPAN

## 參、發明人：(共 2 人)

1. 姓名：(中) 淺野英輔

(英) 浅野英輔

地址：(中) 日本國東京都千代田區丸之内一丁目五番一號新丸大樓日立製作所  
(股) 知的所有權本部內

(英) 日本国東京都千代田区丸の内一丁目5番1号新丸ビル(株)日立製作所 知的所有權本部內

2. 姓名：(中) 新庄廣

(英) 新庄広

地址：(中) 日本國東京都千代田區丸之内一丁目五番一號新丸大樓日立製作所  
(股) 知的所有權本部內

(英) 日本国東京都千代田区丸の内一丁目5番1号新丸ビル(株)日立製作所 知的所有權本部內

## 肆、聲明事項：

◎本案申請前已向下列國家(地區)申請專利  主張國際優先權：

【格式請依：受理國家(地區)；申請日；申請案號數 順序註記】

1. 日本 ; 2003/02/24 ; 2003-045406 有主張優先權

---

(1)

**玖、發明說明****【發明所屬之技術領域】**

本發明係關於處理帳簿傳票之技術，特別是關於作成在進行文字辨識處理時所使用的帳簿傳票定義資料之技術。

**【先前技術】**

在辨識帳簿傳票的金額欄等文字列時，係利用預先登錄在文字辨識要求區域的位置或文字數等資訊的帳簿傳票定義資料，以進行文字辨識。關於此帳簿傳票定義資料作成，至目前為止，有各種手法被提出，那些手法主要係下工夫在使負擔大的定義資料變得更容易之手法。

例如，在記載於日本專利特開 2001-126010 號公報（第 8-9 頁，第 7 圖）之帳簿傳票定義作成方法中，就對於沒有記載資料的帳簿傳票，由格線抽出以及框抽出以自動抽出定義資料；另外，可以辨識預先印刷的文字，藉由與預先登錄的關鍵字辭典對照，設定未被格線圍住的文字寫入框、設定存在於對應關鍵字之位置的框之文字種類；另外，藉由讀取全部的預先印刷文字，不以人手指定框，可自動進行關於全部的定義作成之作業的手法做說明。

**【發明內容】**

(2)

但是，在上述定義作成方法中，在指定框時，在指定框內如沒有預先印刷時，則無法作成文字種類等定義資料。另外，在辨識全面預先印刷文字上，於現況下，由於非常花時間之故，並不實用。另外，對於關鍵字的位置，因帳簿傳票而不同故，不可能對於各種帳簿傳票廣泛地使用關鍵字辭典。

因此，此揭示之主要觀點係用於解決此種問題之技術，不管特定的帳簿傳票是否已經記載完畢，還未記載，即使各種帳簿傳票都可以自動做成帳簿傳票定義資料。

具體為，例如以所指定的讀取區域周邊或者內部的預先印刷文字、記載文字和在定義影像資料輸入時自動抽出的框、格線等佈局資訊為基礎，進行定義資料的自動做成。此定義資料做成方法係即使在讀取區域內部沒有預先印刷，也可以辨識存在於讀取區域周邊的文字列，藉由將該辨識結果轉換為定義資料而可做定義資料的抽出。另外，即使在讀取區域周邊存在多數的文字列時，由文字列的位置或尺寸、框的有無、對於框尺寸之文字列尺寸的比例等，將作為對於讀取區域的關鍵字（以後，稱為項目名稱）的妥當性予以數值化，藉由將最妥當的文字列之辨識結果轉換為定義資料，可以進行定義資料的抽出。

藉由以上的處理，不管記載帳簿傳票／未記載帳簿傳票、對於讀取區域之項目名稱的位置、指定區域內部

(3)

的預先印刷有無，都可以自動做成帳簿傳票定義資料。

另外，可進一步採用各種形態，例如，可以上述的帳簿傳票處理之自動定義資料做成方法構成，也可以使電腦實現此種機能用的電腦程式構成。此處，作為記憶媒體，可以利用軟碟、CD-ROM、DVD、光磁碟、IC卡、IC晶片、ROM卡匣、衝孔紙卡、印刷條碼等符號之印刷物、電腦的內部記憶裝置（RAM或ROM等記憶體）以及外部記憶裝置等電腦可讀取之光學、磁性、電性的種種媒體。上述種種之特徵可以適當地予以組合。

#### 【實施方式】

一面參考圖面一面分成以下的項目來說明合適的一實施形態。

A.系統構成

B.帳簿傳票定義資料的構成

C.帳簿傳票定義資料的作成

C1.項目名稱－定義資料轉換處理

A.系統構成：

第1圖係顯示支援帳簿傳票定義資料的作成之帳簿傳票處理裝置的構造方塊圖。另外，在以下，雖以依據帳簿傳票106的影像資料而重新自動作成帳簿傳票定義資料為例做說明，但是，本帳簿傳票處理裝置也可在已

(4)

經作成的帳簿傳票定義資料新追加登錄別的讀取區域的定義資料。

此帳簿傳票處理裝置係如圖示般，硬體係連接：泛用的個人電腦 101 以及顯示器 102、鍵盤 103、滑鼠 104、掃描器 105 而構成。在個人電腦 101 中安裝有實現帳簿傳票處理裝置之機能用的應用軟體。圖中，顯示作為帳簿傳票處理裝置之機能方塊 107~113。這些機能方塊係藉由上述應用程式構成。當然，也可以硬體構成。

影像輸入部 107 係控制掃描器 105 以達成輸入成為作成帳簿傳票定義資料之樣本的帳簿傳票 106 之影像資料的機能。帳簿傳票定義資料作成部 108 係藉由鍵盤 103 或滑鼠 104 之輸入裝置以指定定義區域，依據此影像資料以達成自動抽出帳簿傳票定義資料的機能。在此時，文字辨識用辭典 110、項目名稱對照用知識辭典 111、項目名稱一定義資訊轉換辭典 112 等各資料庫會被參考。文字辨識用辭典 110 係以各文字單位對照影像資料的形狀和文字用的辭典。項目名稱對照用知識辭典 111 係藉由對照成為文字列和項目名稱的單字，以提升文字辨識率用的辭典。項目名稱一定義資訊轉換辭典 112 係由藉由上述項目名稱對照所獲得的項目名稱，而轉換為成為讀取對象之屬性或文字數等之定義資料用的辭典。

帳簿傳票定義資料輸出部 109 係輸出藉由帳簿傳票定義資料作成部 108 所抽出的定義資料。自動作成的定義資料係登錄在帳簿傳票定義資料用的資料庫 113。

(5)

## B.帳簿傳票定義資料的構造：

第 2 圖係顯示帳簿傳票影像和帳簿傳票定義資料的構造圖。顯示要定在圖的上方之帳簿傳票影像 201 的例子，在下方顯示帳簿傳票定義資料 202 之構造的一例。在帳簿傳票影像 201 中係以左上角為原點，在圖示方向定義為 x、y 軸。

帳簿傳票定義資料 202 的一例可由辨識要求區域座標、框形狀、知識辭典種類、文字數、手寫、鉛字等構成。例如，位於帳簿傳票影像 201 的右上之委託日期的文字辨識定義資料係對應位於定義資料 202 的左上之定義資料。在定義資料中，關於此辨識要求區域，應進行文字辨識處理的矩形範圍係以左上（開始位置）以及右下（結束位置）的各頂點之（x,y）座標所定義。在圖中的例子中，左上的頂點設定為（1200,100）、右下的頂點設定為（1400,150）。另外，框存在故，框形狀設定為「有框」。關於知識辭典種類，由於讀取對象的屬性為日期故，設定為「日期」，和文字數為「12 字」，和文字種類為「鉛字」。

但是，此處所舉之定義資料只為其之一例，此外也可將各種資訊設定作為定義資料。例如，辨識要求區域為鉛字，文字間距為一定等事先知道的狀況時，可將文字間距設定於定義資料，在文字辨識時，藉由使用該資訊，則可以提升辨識率。

(6)

C.帳簿傳票定義資料的作成：

第 3 圖係定義資料自動作成處理的流程圖。為個人電腦 101 的 CPU 對應使用者的指令而實行之處理。處理一開始，則 CPU 先以影像輸入部 107 輸入帳簿傳票的影像資料（步驟 S100），對於帳簿傳票全面實行佈局解析處理（步驟 S200）。即由所輸入的影像資料將表格、框、格線等資訊當成文字資訊予以抽出，同時，被辨識為文字的部份也當成文字資訊予以抽出。

藉由個人電腦 101 的顯示器 102 等之顯示裝置將藉此處理所獲得的佈局資訊提示給使用者。例如，在第 4（a）圖的例子中，在視窗 405 顯示藉由佈局解析所獲得的框抽出結果。在此例中，為了簡化，雖只顯示框抽出結果，但是實際上，也可以依據按鈕或指令等而切換格線、文字行資訊的顯示。

此後要定義的預定區域之格線或框被錯誤抽出時，使用者修正錯誤抽出的佈局資訊（步驟 S300）。修正處理係利用個人電腦 101 的滑鼠 104 等指向裝置來修正顯示在顯示器 102 等的顯示裝置之框或格線等。例如，在第 4（b）圖之例子中，錯誤抽出由佈局解析所獲得的框 406 故，啟動修正按鈕 401，以滑鼠 104 選擇修正框後予以拖曳而修正框（407）。CPU 一檢測到在佈局資訊被施以修正處理時，CPU 依據所修正的資訊，再度進行佈局解析，正確設定關於要定義的預定區域之佈局資訊。

此處理只在定義區域於框或格線等佈局資訊有錯誤

(7)

抽出時進行。因此，在無法確認錯誤抽出時，或在要定義區域以外的部份發生錯誤抽出時，也可跳過此處理故，可以縮短定義資料作成時間。

佈局資訊的修正係藉由對於框或格線等佈局資訊做追加／消除／修正／合併／分割等處理而進行。另外，藉由變更關於在內部所保持的佈局資訊抽出之臨界值，也可以一次修正佈局資訊。例如，變更在內部保持之可抽出的框的最小尺寸、最大尺寸的臨界值，藉由再度進行佈局解析處理，在臨界值變更前無法抽出的框也可一次予以抽出。

藉此處理，一獲得必要的佈局資訊時，接著，CPU進行定義區域設定處理（步驟 S400）。在此處理中，藉由顯示器 102 等的顯示裝置將所獲得之佈局資訊提示給使用者，利用滑鼠 104 等之指向裝置以指定到底是定義哪個區域。此處理可以選擇所獲得的框，在想要定義無框之區域時，也可選擇以滑鼠拖曳等包圍寫有讀取文字列部份的區域。例如，在第 4（d）圖的例子中，要將框 408 當成定義區域時，則使選擇按鈕 402 啟動，藉由以滑鼠 104 選擇框 408 而進行定義區域的設定。

如請使用者指定定義區域時，CPU 則進行佈局資訊一定義資料轉換處理（步驟 S500）。藉由定義區域設定處理，在選擇了當成佈局資訊所被抽出的框時，則由佈局資訊表格取得符合選擇框的資訊，以此為基礎，轉換為定義資料。在定義了無框區域時，則將圍住之框視為

(8)

假想框，作成定義資料。此處所謂定義資料係指藉由辨識要求資訊的矩形座標或框的有無等佈局資訊而可抽出的定義項目。

另外，在定義區域內存在多數框，而且，由各框的縱向、橫向尺寸，如判定全框為 1 文字框時，則可以框數為基礎，進行文字數的設定。例如，在金額欄存在很多藉由位數線所區別的 1 文字框。在定義此區域時，可藉由上述方法以抽出辨識要求區域的矩形座標或框的有無、文字數等的定義資料。

CPU 在此處理前後，也進行項目名稱－定義資料轉換處理（步驟 S600）。詳細處理內容雖之後予以敘述，但是在此處理中，係藉由辨識指定的定義區域之周邊的文字，可以抽出讀取屬性或文字數等定義資料之處理。

藉由這些處理，一獲得定義資料時，接著，CPU 整理定義資料，藉由顯示器等顯示裝置 102，對於使用者提示定義資料。在提示的定義資料有錯誤時，或有未被設定的項目時，使用者修正、追加定義資料（步驟 S700）。在對使用者提示定義資料時，藉由以顏色區分自動設定的項目或未設定項目等，可使使用者容易了解。另外，在自動設定的項目，曖昧性高的項目也同樣予以顏色區分，可敦促使用者注意。此處，雖舉定義資料的使用者提示方法之一例，但是此外，也可考慮各種提示方法。

例如，在第 4（e）圖的例子中，整理由佈局解析以

(9)

及項目名稱所抽出的定義資料予以顯示在視窗 409 中。使用者確認所顯示的定義資料，如全部的定義資料為正確，則不對定義資料加上修正，如錯誤，則使用者修正各定義資料，藉由按下 OK 按鈕 410，結束對於框 408 之定義區域設定。在沒有設定定義區域時，藉由按下清除按鈕 411，則選擇之框的定義區域設定變成無效。

在第 4 (d) 圖之例子中，成為表格形式故，各框的定義資料屬性在列單位具有相同值。例如，位於「銀行名稱」下的框全部為「銀行名稱」之屬性，「分店名稱」也相同。如此，在將定義資料的屬性於列單位相等的區域設定為定義區域時，藉由使用者定義區域的複製機能，可有效率地進行定義作業（步驟 S800）。

例如，在第 5 (f) 圖的例子中，定義區域在定義「銀行名稱」、「分店名稱」、「戶頭號碼」的全區域時，則藉由以上說明的步驟，以設定存在於各項目正下方的區域 412。接著，按下複製按鈕 403 後，如第 5 (g) 圖所示般，藉由以滑鼠 104 拖曳想要複製的區域 413 予以圍住。CPU 對於想要複製的區域 41，印行檢出設定完畢的定義資區域之處理，以及和檢出設定完畢的定義區域 412 其縱向、橫向寬相等的框之處理。在此處理中，對於設定完畢的定義區域 412，在想要複製的區域 413 內，搜尋上下方向，檢出縱向、橫向寬相等的框。接著，CPU 如第 5 (h) 圖所示般，將設定完畢的定義屬性值複製於檢出的框中（414）。此處所謂之定義屬性值係指座

(10)

標資訊以外的文字數或知識辭典種類等定義資料。開始位置或結束位置等座標資訊在各框不同故，這些定義資料係藉由佈局解析所獲得的框資訊予以抽出。

在此例中，雖就對於列之複製做說明，但是對於行之複製也可以同樣實現。另外，此外在檢出縱性、橫向寬相等的框時，藉由顯示器 102 提示給使用者，使用者以滑鼠 104 只選擇想要進行定義屬性複製的框，也可使實行定義屬性複製。

藉由以上的處理，輸出設定的帳簿傳票定義資料（步驟 S900），結束帳簿傳票定義資料的自動作成處理。作成的帳簿傳票如先前說明般，被記憶在帳簿傳票處理裝置，可使用於帳簿傳票的文字辨識。例如，在第 4 圖之例子中，於確認全部的帳簿傳票定義資料被正確設定後，藉由按下保存按鈕 404，可以保存。

#### C1.項目名稱－帳簿傳票定義資料轉換處理：

第 6 圖係項目名稱－定義資料轉換處理 600 的流程圖。在此處理中，對於使用者指定的定義區域，檢出在上方向、左方向鄰接的框（步驟 S601）。此處，CPU 參考預先在帳簿傳票全面抽出的佈局資訊之框資訊表格，檢出相符的框資訊。例如，在第 7 圖的例子中，在將「平成 14 年 12 月 1 日」之區域 705 指定為定義區域時，對於區域 705 之鄰接框係以區域 706 為相符。

接著，CPU 對於存在相符的鄰接框內之文字行，進

(11)

行文字辨識處理(步驟 S602)，進行所獲得之辨識結果是否存在的檢查(步驟 S603)。此處，CPU 參考先前說明的文字辨識用辭典 110，進行取得之光柵圖像和文字之對照。另外，藉由將所獲得的文字列與項目名稱對照用知識辭典 111 對照，實行以確定單字之知識對照處理。

例如，在第 7 圖之例子中，對於指定之定義區域 705 的鄰接框 706 內的文字行 707 參考文字辨識用辭典 110 以及項目名稱對照用知識辭典 111，獲得項目名稱辨識結果「存入指定日期」。所獲得之辨識結果不存在係指無相符之鄰接框時，或者鄰接框內不存在文字行時，另外，文字行雖存在，但是在知識對照失敗。例如，在第 7 圖之例子中，對於區域 701 之鄰接框不存在，只有鄰接的文字行 702 存在。另外，對於區域 703，也不存在鄰接框，在區域 703 的內部存在文字行 704。另外，鄰接框存在有 2 個以上，項目名稱辨識結果存在 2 個以上時，以由文字辨識處理所獲得之可靠度高者為優先。另外，在此情形，夜遊提示給使用者，也可使選擇正確項目名稱。

在可以獲得對於鄰接框內之項目名稱辨識結果時，CPU 將辨識結果之項目名稱轉換為定義資料(步驟 S609)。在此處理中，藉由參考先前說明的項目名稱－定義資訊轉換辭典 111，轉換為對於項目名稱之定義資料。第 8 圖係顯示項目名稱－定義資訊轉換辭典 111 之一例。如以第 7 圖之區域 706 內的「存入指定日期」為例做說明

(12)

時，此項目名稱係存在於項目名稱－定義資訊轉換辭典 111，與此項目名稱相符的知識辭典種類為「日期」、文字數為「12 字」。如此，由項目名稱抽出定義資料。另外，附隨項目名稱之定義資料可不管知識辭典種類或文字數而設定各種資訊。例如，在知識辭典種類或文字數以外，也可考慮文字種類等。

在步驟 603 中，在無法獲得對於鄰接框之項目名稱辨識結果時，則進行指定之定義區域內的文字行抽出處理。此處，CPU 參考預先在帳簿傳票全面所抽出的佈局資訊的文字行資訊表格，檢出存在於指定區域內的文字行資訊。對於抽出的文字行，進行文字辨識處理（步驟 S604），進行所獲得的辨識結果是否存在的檢查（步驟 S605）。此處，CPU 與上述相同，利用文字辨識用辭典 110、項目名稱對照用知識辭典 111，進行文字辨識。

例如，在第 7 圖之例子中，於將「委託日 年 月」之區域 703 指定為定義區域時，將存在於指定之定義區域 703 內部的文字行 704 參考文字辨識用辭典 110 以及項目名稱對照知識辭典 111，獲得項目名稱辨識結果「委託日」。在可以獲得對於內部文字行之項目名稱辨識結果時，CPU 將辨識結果之項目名稱轉換為定義資料（步驟 S609）。

在無法獲得對於內部文字行之項目名稱辨識結果時，對於指定的定義區域，檢出與上方向、左方向鄰接之文字行（步驟 S606）。此處，CPU 參考預先在帳簿傳票

(13)

全面所抽出的佈局資訊的框資訊表格，檢出相符之文字行資訊。例如，在第 7 圖之例子中，在將「——殿」之區域 701 指定為定義區域時，對於區域 701 之鄰接文字行係相當於區域 702。

接著，CPU 對於相符之鄰接文字行，進行文字辨識處理（步驟 S607），進行所獲得之辨識結果是否存在之檢查（步驟 S608）。此處，CPU 與上述相同，利用文字辨識用辭典 110、項目名稱對照知識辭典 111 以進行文字辨識。例如，在第 7 圖之例子中，將對於指定之定義區域 701 的鄰接文字行 702 參考文字辨識用辭典 110 以及項目名稱對照知識辭典 111，獲得項目名稱辨識結果「委託人」。

在可獲得對於鄰接文字行之項目名稱辨識結果時，CPU 將辨識結果的項目名稱轉換為定義資料（步驟 S609）。在無法獲得對於鄰接文字行之項目名稱辨識結果時，將指定之定義區域視為不具有項目名稱之區域，知識辭典種類或文字數等之定義資料當成未設定而結束。

CPU 對於全部指定定義區域進行以上處理。另外，在此次的項目名稱抽出處理中，雖以鄰接框內文字行、指定定義區域內文字行、鄰接文字行之順序設定優先度，但是，也可依據帳簿傳票種類而變更優先度。另外，不使用 3 個文字行，例如也可以只使用鄰接框內文字行。藉由如此，在項目名稱出現的位置受到限制之帳簿傳票等，可以進行更正確之項目名稱抽出定義資料作成。

(14)

如以上說明般，依據揭示之技術，可以儘可能使帳簿傳票定義資料的作成自動化，關於無法自動化的處理，藉由部份地人工之介入，可以更圓滑地支援帳簿傳票定義資料作成。

使用者在修正、追加定義資料時，藉由在顯示裝置以顏色區分自動設定的項目或未設定項目等，可以使之成為使用者容易了解的畫面構造。另外，在自動設定的項目中，曖昧性高的項目也可同樣予以顏色區分，可促使使用者注意。

另外，不管特定帳簿傳票或者記載完畢、未記載，任何種類之帳簿傳票都可以自動作成帳簿傳票定義資料。

以上，所揭示的技術並不限定於實施形態，在不脫離其旨趣之範圍內，不用說可以有種種之構造。例如，以上的控制處理在以軟體實現之外，也可以硬體實現。另外，也可以帳簿傳票處理裝置的文字辨識手段實施帳簿傳票定義資料的作成來構成。

如依據所揭示的技術，不管特定的帳簿傳票或者記載完畢、未記載，任何種類之帳簿傳票都可以自動作成帳簿傳票定義資料。

#### 【圖式簡單說明】

第 1 圖係帳簿傳票處理裝置的概略構造圖。

第 2 圖係顯示帳簿傳票影像以及帳簿傳票定義資料

(15)

的構造圖。

第 3 圖係帳簿傳票定義資料自動做成處理的流程圖。

第 4 圖係顯示說明帳簿傳票定義資料的作效用之顯示示例圖。

第 5 圖係顯示說明帳簿傳票定義資料作效用之顯示示例圖。

第 6 圖係帳簿傳票定義資料作成之項目名稱一定義資料轉換處理的流程圖。

第 7 圖係顯示帳簿傳票定義資料之對於所指定的定義區域的項目名稱的位置圖。

第 8 圖係顯示帳簿傳票定義資料作成的項目名稱一定義資料轉換辭典的一例圖。

#### 主要元件對照表

- 101 個人電腦
- 102 顯示器
- 103 鍵盤
- 104 滑鼠
- 105 掃描器
- 106 帳簿傳票
- 107 影像輸入部
- 108 帳簿傳票定義資料作成部
- 109 帳簿傳票定義資料輸出部

(16)

- 110 文字辨識用辭典
- 111 項目名稱對照用知識辭典
- 112 項目名稱-定義資訊轉換辭典
- 113 帳簿傳票定義資料
- 201 帳簿傳票影像
- 202 帳簿傳票定義資料

### 伍、中文發明摘要

發明之名稱：帳簿傳票之定義資料作成方法及帳簿傳票之處理裝置

減輕帳簿傳票定義資料作成時的使用者負擔。帳簿傳票處理裝置係對於使用者指定的區域，進行佈局資訊的抽出（步驟 200）以及與所指定的區域有關連的項目名稱的抽出，將這些資訊轉換為定義資料（步驟 500、步驟 600），藉此以實施帳簿傳票定義資料的自動作成。藉此，只指定定義資料，可以自動作成定義資料故，能夠減輕使用者的帳簿傳票定義資料的設定負擔。

### 陸、英文發明摘要

發明之名稱：

(1)

### 拾、申請專利範圍

1.一種帳簿傳票定義資料作成方法，其特徵為：

取得帳簿傳票的影像資料；

由該影像資料抽出文字資訊的佈局資訊；

由對應所指定的定義區域之上述佈局資訊，抽出關於該定義區域的位置之第 1 定義資料；

辨識存在於前述定義區域的周邊或者內部的文字資訊；

將辨識結果轉換為關於該定義區域的屬性之第 2 定義資料。

2.如申請專利範圍第 1 項所記載之帳簿傳票定義資料作成方法，其中：在前述定義區域的附近檢查前述文字資訊的存在；

於檢查結果，於前述定義區域的附近沒有檢出文字資訊的存在的場合，在該定義區域的內部檢查文字資訊的存在；

於檢查結果，於前述定義區域的內部也沒有檢出文字資訊的存在的場合，檢查位於該定義區域的上方向以及左方向之文字資訊的存在。

3.如申請專利範圍第 1 項所記載之帳簿傳票定義資料作成方法，其中：於前述定義區域於列方向連續的場合，由對應各定義區域的前述佈局資訊抽出關於該各定義區域的位置之第 1 定義資料，前述第 2 定義資料是把關於前述各定義區域的屬性當成第 2 定義資料來加以複製。

(2)

4.如申請專利範圍第 1 項所記載之帳簿傳票定義資料作成方法，其中：於前述佈局資訊錯誤的場合，以經過修正的資訊為基礎，再度抽出佈局資訊。

5.如申請專利範圍第 1 項所記載之帳簿傳票定義資料作成方法，其中：經由自對應到前述定義區域之前述佈局資訊以求得該定義區域內的各框之縱橫比來判定文字框的有無，於判定為文字框的場合，計算文字框數以抽出文字數的定義資料。

6.一種帳簿傳票處理裝置，是針對依據帳簿傳票的影像資料以進行所記載的內容之文字辨識處理時所使用的定義資料的作成，其特徵為具有：

取得帳簿傳票的影像資料之手段；和

於前述影像資料，抽出框、格線以及文字行等佈局解析資訊之手段；和

由對應所指定的定義區域之前述佈局解析資訊來抽出關於該定義區域的位置的定義資料之手段；和

由存在於前述定義區域周邊或者內部的框以及文字行來抽出該定義區域的項目名稱之手段；和

進行前述項目名稱的文字辨識之手段；和

將由前述文字辨識處理所獲得的辨識結果與項目名稱辭典對照之手段；和

將由前述對照結果所獲得的項目名稱轉換為顯示該定義區域的屬性的定義資料之手段；和

整理前述定義資料以輸出於帳簿傳票定義資料檔案

(3)

之手段。

7.如申請專利範圍第 6 項所記載之帳簿傳票處理裝置，其中：具有，於上述佈局解析資訊錯誤的場合，以經過修正的佈局解析修正資訊為基礎，經由再度進行佈局解析處理，來修正格線或框等之佈局解析資訊之手段。

8.一種帳簿傳票處理裝置；是具有，讀取帳簿傳票以取得影像資料之影像輸入手段，和及文字辨識來自該影像輸入手段的影像資料之文字辨識手段，其特徵為：

前述文字辨識手段係由來自前述影像輸入手段的影像資料抽出文字資訊的佈局資訊，由對應所指定的定義區域之前述佈局資訊來抽出關於該定義區域的位置之第 1 定義資料，辨識存在於前述定義區域的周邊或者內部之文字資訊，將辨識結果轉換為關於該定義區域的屬性之第 2 定義資料，彙整保存該第 2 定義資料和上述第 1 定義資料。

9.如申請專利範圍第 8 項所記載之帳簿傳票處理裝置，其中：於前述定義區域於列方向連續的場合，前述文字辨識手段由對應前述各定義區域之前述佈局資訊抽出關於該各定義區域的位置之第 1 定義資料，前述第 2 定義資料是把關於前述各定義區域的屬性當成第 2 定義資料來加以複製。

10.如申請專利範圍第 8 項所記載之帳簿傳票處理裝置，其中：前述文字辨識手段經由自對應到前述定義區域之前述佈局資訊以求得該定義區域內的各框之縱橫比來

(4)

判定文字框的有無，於判定為文字框的場合，計算文字框數以抽出文字數的定義資料。

圖 1

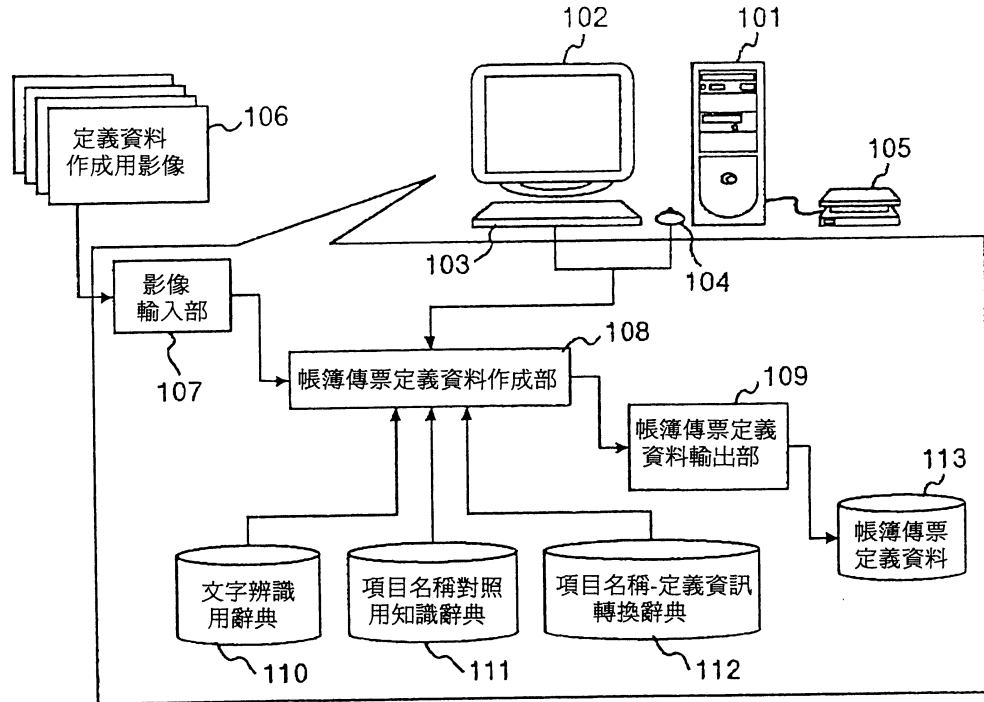


圖2

綜合存入委託書

委託人 \_\_\_\_\_ 委託日 月 日

先生/小姐 \_\_\_\_\_ 印章 \_\_\_\_\_

存入指定日期	店號碼	顧客號碼	銀行代碼
平成14年12月 1日	001	0000001	000001

銀行名稱	分行名稱	戶頭號碼	領取人	金額	存入區分	收續費
A銀行	A分行	1234567	A田B郎	80000		420
B銀行	B分行	2345678	C山D子	70000		420
C銀行	C分行	3456789	F口G人	60000		420
D銀行	D分行	4567890	H川I美	50000		420
E銀行	E分行	5678901	J野K男	40000		420
F銀行	F分行	6789012	L本M香	30000		420
G銀行	G分行	7890123	N木O雄	20000		210
H銀行	H分行	8901234	P藤Q穂	10000		210
小計				360000		2940
合計金額				362940		

201



定義資料辨識要求區域座標		框形狀	知識辭典種類	文字數	手寫、鉛字
開始位置	結束位置				
(1200,100)	(1400,150)	有框	日期	12	手寫
(100,100)	(300,150)	無框	姓名	30	手寫
(100,200)	(250,250)	有框	日期	12	鉛字
(800,300)	(1000,350)	有框	金額	15	鉛字
(800,700)	(1000,750)	有框	金額	15	鉛字
⋮	⋮	⋮	⋮	⋮	⋮

202

圖 3

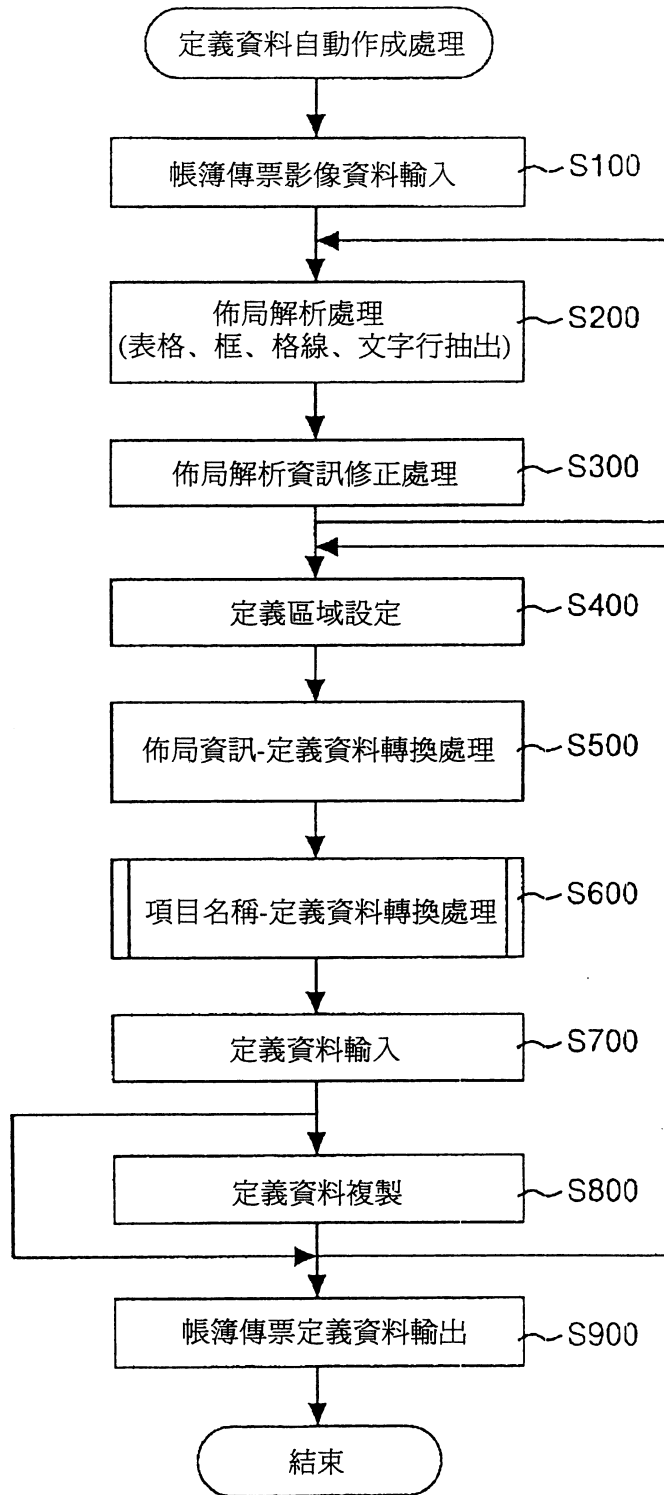


圖 4

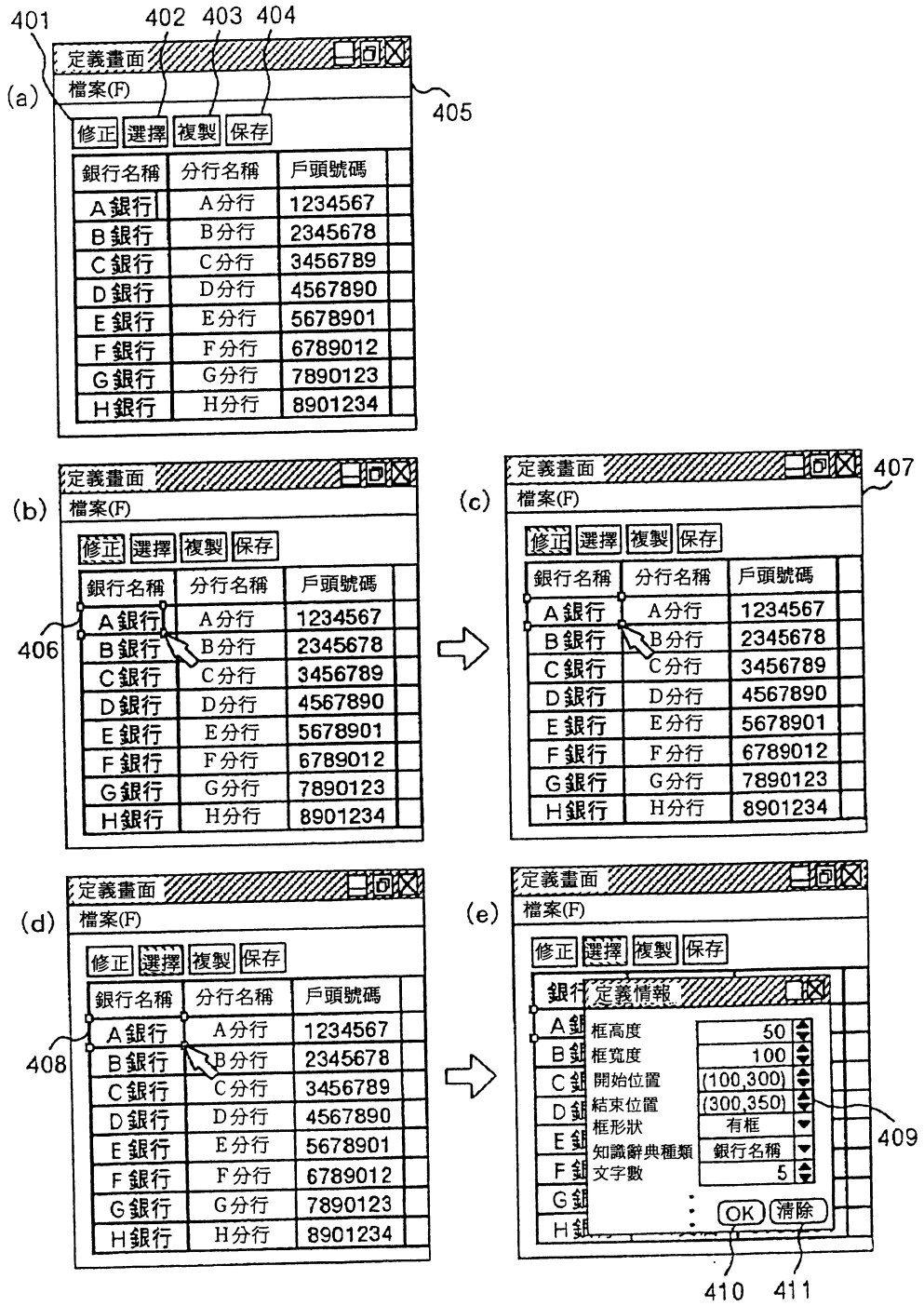




圖6

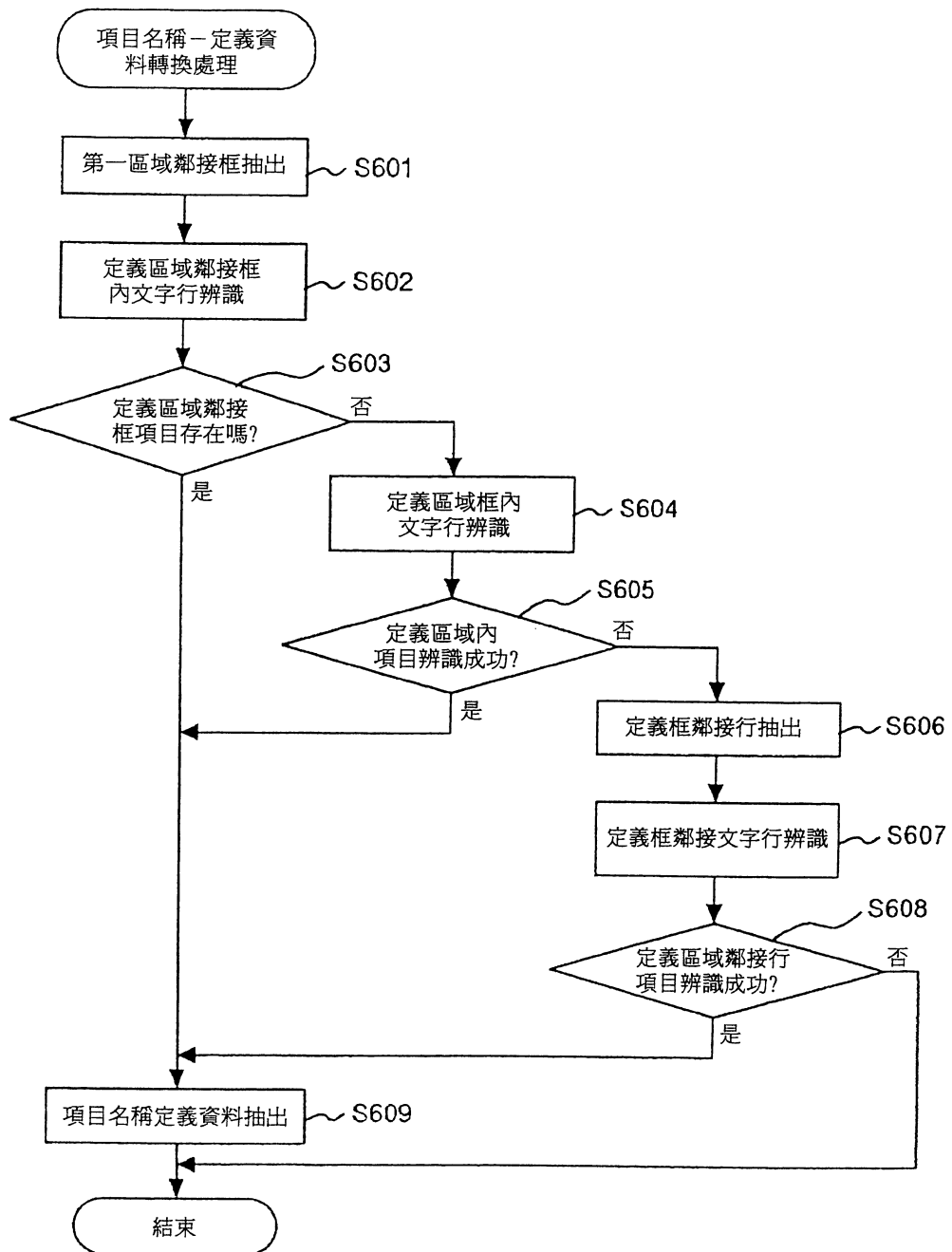


圖 7

702

委託人

701 先生/小姐

704

703 委託日 月 日

706

707

存入指定日期	店號	顧客號碼	銀行代碼
平成 14年12月 1日	001	0000001	000001

705

709 710

銀行名稱	分行名稱	戶頭號碼	領取人	金額
A 銀行	A 分行	1234567	A 田 B 郎	80000
B 銀行	B 分行	2345678	C 山 D 子	70000
C 銀行	C 分行	3456789	F 口 G 人	60000
D 銀行	D 分行	4567890	H 川 I 美	50000
E 銀行	E 分行	5678901	J 野 K 男	40000
F 銀行	F 分行	6789012	L 本 M 香	30000
G 銀行	G 分行	7890123	N 木 O 雄	20000
H 銀行	H 分行	8901234	P 藤 Q 穂	10000
			小計	360000
			合計金額	362940

708

713

712

711

圖 8

項目名稱	定義資訊	
	知識辭典種類	文字數
姓名	姓名	30
寄出人	姓名	30
⋮	⋮	⋮
地方	住所	64
本籍	住所	64
⋮	⋮	⋮
存入金額	金額	15
總額	金額	15
⋮	⋮	⋮
TEL	電話號碼	11
行動電話	電話號碼	11
⋮	⋮	⋮
存入指定日期	日期	12
日期	日期	12
⋮	⋮	⋮

~111

柒、(一)、本案指定代表圖為：第 1 圖

(二)、本代表圖之元件代表符號簡單說明：

101：個人電腦，102：顯示器，103：鍵盤，104：滑鼠，105：掃描器，106：定義資料作成用影像，107：影像輸入部，108：帳簿傳票定義資料作成部，109：帳簿傳票定義資料輸出部，110：文字辨識用辭典，111：項目名稱對照用知識辭典，112：項目名稱－定義資訊轉換辭典，113：帳簿傳票定義資料。

捌、本案若有化學式時，請揭示最能顯示發明特徵的化學式：