

公告本

附件

第 84101589

號專利申請案中文說明書修正本

(含申請專利範圍)

民國 85 年 5 月 5 日

389865

389865

申請日期	84 年 2 月 21 日
案 號	84101589
類 別	G06F 1/00, G06T 3/00

Int. Cl⁶

A4
C4

修正
年 月 日
85. 5. 28 補充

(以上各欄由本局填註)

發明專利說明書

一、發明 名稱	中文使用新穎之由果溯因機率方式來對輸入表象作自動詮釋之系統與方法，以及最佳訓練之資訊處理網路 英文 System and method for automated interpretation of input expressions using novel a posteriori probability measures and optimally trained information processing network
二、發明 人	姓 名 (1) 克理斯多夫·柏格斯 Burges, Christopher John (2) 約翰·丹克 Denker, John S. 國 籍 (1) 美國 (2) 美國 (1) 美國新澤西州·費弗德·安朵拉區十一號 11 Andorra Terrace, Freehold, NJ 07728, USA 住、居所 (2) 美國新澤西州里歐納度庫斯曼道六號 6 Koosman Drive, Leonardo, NJ 07737, USA
三、申請人	姓 名 (1) 美國電話電報股份有限公司 (名稱) AT&T Corp. 國 籍 (1) 美國 (1) 美國紐約州·紐約市美州大道三十二號 32 Avenue of the Americas, New York, NY 10013-2412, U.S.A. 住、居所 (事務所) 代 表 人 姓 名 (1) 皮·汪爾德 Wilde, P. V. D.

裝

訂

線

公告本

附件

第 84101589

號專利申請案中文說明書修正本

(含申請專利範圍)

民國 85 年 5 月 5 日

389865

389865

申請日期	84 年 2 月 21 日
案 號	84101589
類 別	G06F 1/00, G06T 3/00

Int. Cl⁶

A4
C4

修正
年 月 日
85. 5. 28 補充

(以上各欄由本局填註)

發明專利說明書

一、發明 名稱	中文使用新穎之由果溯因機率方式來對輸入表象作自動詮釋之系統與方法，以及最佳訓練之資訊處理網路 英文 System and method for automated interpretation of input expressions using novel a posteriori probability measures and optimally trained information processing network
二、發明 人	姓名 (1) 克理斯多夫·柏格斯 Burges, Christopher John (2) 約翰·丹克 Denker, John S. 國籍 (1) 美國 (2) 美國 住、居所 (1) 美國新澤西州·費弗德·安朵拉區十一號 11 Andorra Terrace, Freehold, NJ 07728, USA (2) 美國新澤西州里歐納度庫斯曼道六號 6 Koosman Drive, Leonardo, NJ 07737, USA
三、申請人	姓名 (名稱) (1) 美國電話電報股份有限公司 AT&T Corp. 國籍 (1) 美國 住、居所 (事務所) (1) 美國紐約州·紐約市美州大道三十二號 32 Avenue of the Americas, New York, NY 10013-2412, U.S.A. 代表人姓名 (1) 皮·汪爾德 Wilde, P. V. D.

裝

訂

線

(由本局填寫)

承辦人代碼：
大類：
I P C 分類：

A6

B6

本案已向：

國(地區) 申請專利，申請日期： 案號：
美國 1994 年 8 月 4 日 08/286147

，☐有 ☐無主張優先權
☒無主張優先權

有關微生物已寄存於：

，寄存日期：

，寄存號碼：

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

五、發明說明(1)

相關專利申請案

本案為申請人於1991年12月31日提申之名為「用以在影像識別系統中，自動區分及辨識的圖形系統」申請序號No. 07/816,415，以及名為「阿拉伯數字的影像區分機制」申請序號07/816,414，兩案之部分接續案，並在此作為參考文件。

發明領域：

本發明大體而言係關於一種用以自動詮釋輸入表象的方法和系統，此表象可能是手寫的字符，使用的是一種新式的由果溯因機率量度以及最佳化訓練之神經資訊處理網路。

習知技術之簡述：

目前，在製造能正確地將記錄在各種媒體中，可能是相連的阿拉伯數字字符串作詮釋（亦即辨識）的機具上存在有極大的商業利益。例如，美國的郵遞系統在其全國郵件分送作業中，很快地將廣泛使用該類機具，以期能正確辨識出郵件上手寫的郵遞區號。

目前，已有許多習知的字符辨認系統被開發出來，且應用於各種環境。許多此等系統及相關技術被描述於下列的技術文獻中：

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

泉

五、發明說明 ()

publications:

- Y. Le Cun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel, "Handwritten Digit Recognition with a Back-Propagation Network", pp. 396-404 in *Advances in Neural Information Processing 2*, David Touretzky, ed., Morgan Kaufman (1990);
- J.S. Bridle, "Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition", in *Neuro-computing: Algorithms, Architectures and Applications*, F. Fogelman and J. Héroult, ed., Springer-Verlag (1989);
- J.S. Bridle, "Training Stochastic Model Recognition Algorithms as Networks Can Lead To Maximum Mutual Information Estimation of Parameters", in *Advances in Neural Information Processing 2*, David Touretzky, ed., Morgan Kaufman (1990);
- O. Matan, J. Bromley, C.J.C. Burges, J.S. Denker, L.D. Jackel, Y. LeCun, E.P.D. Pednault, W.D. Satterfield, C.E. Stenard, and T.J. Thompson, "Reading Handwritten Digits: A ZIP code Recognition System", *IEEE Computer* 25(7) 59-63 (July 1992);
- C.J.C. Burges, O. Matan, Y. LeCun, J.S. Denker, L.D. Jackel, C.E. Stenard, C.R. Nohl, J.I. Ben, "Shortest Path Segmentation: A Method for Training a Neural Network to Recognize Character Strings", *IJCNN Conference Proceedings 3*, pp. 165-172 (June 1992);
- C.J. Burges, O. Matan, J. Bromley, C.E. Stenard, "Rapid Segmentation and Classification of Handwritten Postal Delivery Addresses using Neural Network Technology", Interim Report, Task Order Number 104230-90-C-2456, USPS Reference Library, Washington D.C. (August 1991);
- Edwin P.D. Pednault, "A Hidden Markov Model For Resolving Segmentation and Interpretation Ambiguities in

經濟部中央標準局員工消費合作社印製

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(3)

Unconstrained Handwriting Recognition", Ball Labs Technical Memorandum 11352-090929-01TM, (1992); and

- Ofer Matan, Christopher J.C. Burges, Yann LeCun and John S. Denker, "Multi-Digit Recognition Using a Space Displacement Neural Network", in Neural Information Processing Systems 4, J. M. Moody S.J. Hanson and R.P. Lippman, eds., Morgan Kaufmann (1992).

雖然以上提及的習知系統各有所不同，而所具有共同的特徵是結構化和功能的特徵。

特別是，各習知系統至少取得一個影像 I，影像 I 是一串將被此系統加以詮釋之可能相連的字符。一般而言，針對一個已知的字母，系統必須自其中選取「最佳」詮釋的可能詮釋數目等於可使用字母中之字符及適用之字型限定被連接在一起之可能的字符串數目。對於郵遞區號碼辨識應用而言，各種被允許的詮釋被郵遞區號碼的長度加以限定，亦即，它必須是 5 或 9 個數碼。

根據傳統，取得的字串影像基本上先予以前置處理以去除底線，空間雜訊等。處理過的影像 I 則被切割或區分為可方便管理的子影像。存在於各組相鄰切割線之間的子影像被稱為「細元 (cells)」。在某些狀況下，在兩細元間的邊界被確定是為一個「限定切割」，確切地落於兩個字符之間，在其他的狀況下，所謂的切割是非限定的，而此切割是否落於二字符之間並未即刻確定，直到進一步處理被實施為止。之後相鄰的影像細元結合起來以組成影像「分段 (Segments)」。然後，這些影像分段由左至右

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(4)

依序構成可接受的影像「串接分段」其實際上包含在前置處理過後之影像中所有的像素(pixel)。特別的是，一種非循環指向圖形用來建構一種可接受的影像「串接分段」模型。一般而言，此模型是藉由將各個影像分段與一個存在於指向非循環圖形中的節點相結合而被建構的。在此圖形中的節點然後與指向的弧線相連。大體而言，圖形中的兩個節點僅有在它們所代表的影像分段在一個可接受影像串接分段中是合法相鄰的情況下才被相連。

當圖形完全建構好之後，其間通過的每一條路徑相對應於前置處理後之影像中的一個影像串接分段，且每一個可能的影像串接分段對應一個通過圖形的特定的路徑。在圖形已建立之後，使用遞迴修剪技術以從圖形中去除任何對應於某一特定影像分段的任何節點，所謂特定影像分段是在經前置處理後之影像中，橫跨在一限定切割線上的影像分段。在圖形修剪完之後，與仍留在該被修剪的圖形中之一節點相連之各個影像分段被傳送至一個神經網路辨識器用以進行分類和給分(scoring)。根據此分類和給分，修剪圖形中的各節點被指定一個「分數」，此分數是導源自於指定給相關影像分段的辨識器分數。基本上，各個辨識器分數藉由一個將辨識器分數正規化之計算程序被轉換為機率值。之後，針對修剪圖形中的各條路徑，計算出一個路徑分數(亦即聯合機率)，所使用的方法單純地是將指定給路徑上各節點的「分數」相乘。根據這種多字符辨識(MCR)機制，在修剪圖形中的最高給分路徑對應

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

五、發明說明(5)

於取得之影像中“最佳”的影像串連分段和字串詮釋。這種技術可詳見以下的兩篇專利申請案，1991年12月31日申請之序號No. 07/816,414“阿拉伯數字影像串接分段機制”，以及1991年12月31日申請之序號No. 07/816,415“使用於影像辨識系統中自動串接分段及辨識的繪圖系統”，以上二篇專利申請案均作為本發明之參考文件。

雖然習知的方法已適用於商業或實驗性的字符辨識系統，此種系統的性能仍不夠完善，特別是在要求高和即時處理的應用中。尤其是，習知技術中的MCR系統一般而言是藉著僅辨識一個支援某一已知詮釋的串接分段來操作的。這種方法運用的前提是僅存在一個“最佳”的串接分段。根據這種習知的方法，此一個“最佳”串接分段的分數是在辨識程序中唯一被考慮的分數。因此，習知技術的方法相當於錯誤地假設正確的影像串接分段是已知的。根據這種假設，各別的字符分數將予以正規化，以針對允許碼和字母中的特定字符，計算其機率。這將無可挽回地丟棄某些有用的資訊，而這些有用的資訊是與針對影像中特定分段，分段演算法執行性能的良劣程度有關。根據此種假設所設計的習知技術MCR系統通常被稱之為“最大可能性序列預估”(MLSE)機。

除了選取一個影像的詮釋，若干習知MCR系統亦通常提供一個分數，意即提供若干機率的指示數值，亦即詮釋正確的機率。在許多應用中，我們希望取得一個能代表

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

象

五、發明說明(6)

正確機率值的分數，以便於將M C R系統得到的結果和其他資訊來源結合起來。然而，習知技術中的M C R系統強調的是選取「最佳」的詮釋，而未強調正確地給分。因此，得到的分數通常包含規律性的錯誤，而這些錯誤的級數頗大。

因此有必要針對各種媒體中描述之符號序列的詮釋，設計出一種改良的方法和系統。

本發明的目的

本發明的整體目的在於提出一種改良的方法和系統，以詮釋輸入表象，這些輸入表象可以是利用印刷或手寫草體方式表現或記錄在媒體中的字符串。

本發明的另一個目的是提供一種自動詮釋字符串的系統和方法，它是使用一種由果溯因的機率值來選取最佳的字符串詮釋。

本發明的另一個目的是提供一種自動詮釋字符串的系統和方法，藉著由果溯因機率值與已知實例像素影像的結合，可歸納性地導出各個由果溯因機率值。

本發明的另一個目的是提供一種自動詮釋字符串的系統和方法，它可以詮釋任意長度的字符串，且可輕易地搭配類似具有自動文句詮釋功能的系統。

本發明的另一個目的是提供一種多字符辨識系統，其中，一種最佳訓練神經計算網路叢在進行影像串接分段及字符串詮釋程序時被使用。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(7)

本發明的另一個目的是提供一種多字符辨識的方法，其中的影像串接分段和字元串詮釋被結合成單一的適應學習程序，這個程序係以最佳化訓練的神經計算網路叢來加以實施，用以將正確字符串詮釋的給分最大化。

本發明的另一個目的是提供一種使用新式資料結構的系統，而此資料結構是根據一種特別改良的非循環指向圖形構成，其中通過該圖形的各條路徑同時代表著一個影像串接分段和一個字符串詮釋。

本發明的另一個目的是將分數指定給選出可能之影像詮釋，特別是，能針對這些選出的詮釋，代表其正確機率預估值的分數。

本發明的另一個目的是提供一種系統，其中指定給各特定字符串詮釋的由果溯因機率是以一個比值來定義，而此比值的分子是將通過代表相同字符詮釋的圖形之所有路徑的路徑分數相加起來以計算得出，而上述比值的分母是將通過代表具有相同的字符數目之所有可能的字符串詮釋之所有路徑的路徑分數相加求得，。

本發明的另一個目的是提供一種能設計成攜帶式裝置的多字符手寫辨識系統。

本發明另一個目的是提供一種字符串詮釋的方法，其中，使用Viterbi演算法來辨識出那一個字符串詮釋具有最佳路徑分數；其中的前向演算法(Forward Algorithm)接著被用來計算出代表利用Viterbi演算法辨認出來的字符串詮釋之所有路徑分數的真實總和值，且其中的前向

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

五、發明說明(8)

演算法是用來計算出一個針對上述正確計算求得之總和的正規化常數，計算的方式係藉由將通過代表所有可能的字符串詮釋之圖形的所有路徑分數相加起來。

本發明的另一個目的是提供一種字符串詮釋方法，其中使用了射線搜尋演算法 (Beam Search Algorithm) 來辨認多個具有最佳路徑分數集合之競爭字符串詮釋；其中的前向演算法被用來對各個字符串詮釋，計算出所有路徑分數的正確總和值，而這些路徑分數代表著使用 Viterbi 演算法辨認出來的競爭字符串詮釋；且藉著將通過代表所有可能字符串詮釋之圖形的所有路徑分數相加起來，前向演算法則被用以針對各個競爭字符串詮釋計算出單一正規化常數。

本發明的另一個目的是提供一種具有操作學習模式的輸入表象詮釋系統，其中的圖形和神經告知處理網路叢兩者皆在一個或多個訓練階段期間藉由最佳化地調整神經網路之參數，被用以訓練該系統。

本發明的另一個目的是提供一種系統，其中在神經網路訓練期間將進行靈敏度分析，而在神經網路中各個可調參數調整的方向是能夠增加可知正確字符串詮釋的由果溯因機率，以及降低可知不正確詮釋的由果溯因機率。

本發明的另一個目的是提供一種輸入表象詮釋系統，在此系統進行學習操作模式期間，使用了 Baum-Welch 演算法來計算整體系統變動時所產生分數變化的靈敏度，其中所稱之系統變動乃肇因於神經網路中各個可調參數的遞

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(9)

增變量。

以上提及或其他的目的可在以下說明和申請專利範圍中清楚得知。

本發明總結：

根據本發明之第一個態樣，提出一種用以詮釋表現在一媒體中之輸入表象的方法和系統，表現的方式可能是印刷或是手書寫。

一般而言，此系統取得代表輸入表象的輸入資料集合。取得的輸入資料集合然後被分割為一組分段，這些分段用以指定一組串接分段。本系統接著使用一種新式的資料結構針對輸入表象隱性地表示出各串接分段和各個可能的詮釋。這個資料結構可以具有二次元節點陣列的一個指向非循環圖形來表示，這些節點以行列的形式排列且以指向的弧線作選擇性地連接。通過節點且沿著指向弧線延伸的各條路徑代表著輸入表象的一個串接分段和一個可能的詮釋。輸入表象的所有串接分段和所有可能的詮釋是由一通過圖形的路徑組所隱性的表示。

對於圖形中各列的節點而言，將針對已知的符號集合，使用一個最佳化訓練神經資訊處理網路來產生一組分數。配合該圖形，這些分數隱性地針對圖形中通過的各條路徑，指定一個路徑分數。本系統利用這些路徑分數來辨識出最佳的符號序列詮釋且計算出相關的由果溯因機率。藉著分析整個取得的輸入資料集合，可求得各個由果溯因機

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

東

五、發明說明(10)

率，進而得到各個符號序列詮釋相關的一個高可靠度機率。本發明的原理可藉由視覺表現之任何可符號序列予以實施，包括任意長度的手寫字符串。本系統亦可輕易地與自動文句詮釋系統搭配使用。

在本發明的第一個實施例中涉及字符串辨識，該系統確定出具有圖形中最高給分路徑的字符串詮釋。為了確定這個詮釋是否可靠，此系統亦產生一個針對此字符串詮釋的由果溯因機率作為輸出。計算得出的機率值是一個具有分子和分母的比值。分子等於通過代表一已知的字符串詮釋圖形中所有路徑的路徑分數總和。分母等於通過代表了所有可能的字符串詮釋圖形中所有路徑的路徑分數總和。若此機率值小於一預設的限值，使用者無法保證該詮釋是可靠的，則使用者被告知在進一步動作前須完成其他步驟。

在本發明的第2個實施例中，此系統先找出通過具有一組最高路徑得分之圖形的一組路徑。對於此組中的各條路徑，系統將辨識出對應的字符串詮釋，並且求出該詮釋之一個由果溯因機率值（包含取自其他具有相同貢獻之路徑的貢獻）。此系統辨識出由該被找出之路徑組所代表之可能的字符串詮釋集合。然後這組可能字符串詮釋的由果溯因機率可被算出。此系統分析計算所得到的該組由果溯因機率，以便確定那一個可能字符串詮釋具有最大的由果溯因機率。根據該項分析，此系統將產生兩個輸出，（i）一個或多個具有高由果溯因機率的字符串詮釋，（i i

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明(11)

) 各個字符串詮釋的由果溯因機率的一個正確預估值。

在第2個實施例中，針對各個競爭字符串詮釋所計算求得的由果溯因機率是一個具有分子和分母的比值。分子部份等於通過代表了競爭字符串詮釋之圖形中所有路徑的路徑得分總和。分母部份等於通過代表所有可能的字符串詮釋之圖形中所有路徑的得分總和。

根據本發明的第2個態樣，一種新式的方法和裝置被提供給本發明之符號序列詮釋作最佳化的訓練。這是藉由提供該系統一獨特的操作學習模式而被達成。

在操作學習模式進行時，此系統將處理大批代表已知輸入表象的訓練影像。對於各個被處理的訓練影像而言，系統遞增性地調整該組能影響各神經網路的參數。各個遞增性調整的方向是使得已知正確字符串詮釋的平均機率能夠遞增，而使得已知不正確符號序列詮釋的平均機率能夠降低。

本發明中的系統和方法可適用於任何視覺表象的字符串詮釋，包括記錄在電氣被動式媒體，例如紙張，塑膠，織品的圖形，或者記錄在電氣主動式媒體，例如壓敏手寫平面，以及「觸動螢幕」手寫與顯示平面，這些媒體均是業界所習知的。

爲了更完整地了解本發明的目的，以下實施例的詳細說明應參照所附的圖式一起被閱讀，其中：

圖1爲一顯示被使用來實施本發明實施例之字符串詮釋系統之各種組件之系統方塊圖；

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

泉

五、發明說明(12)

圖 2 為一本發明之字符串詮釋系統的方塊圖；

圖 3 為一以手書寫方式寫下之郵遞區號的一個前置處理影像；



圖 4 A 至 4 C 是一組圖 3 中所示郵遞區號的前置處理影像，每一個圖均顯示出一組添加的切割線，這些切割線是在本發明之字符串詮釋方法中的影像細元產生階段進行時所生成的；

圖 5 是形成在圖 4 中之切割線之間之影像「細元」（亦即子影像）的列表；

圖 6 為結合圖 4 及 5 中所示相鄰影像細元而形成的影像「分段」之一列表；

圖 7 為一列出由圖 6 中之空間上連續的影像分段之相連集合所形成之許多合法的「影像分段」中的三個的表列；

圖 8 是本發明中新式資料結構的圖形表示，藉此圖形可用以表示影像分段，利用影像分段將構成可能的影像串接分段，可能的字符串詮釋以及指定給可能字符串詮釋的分數；

圖 9 是本發明中字符串詮釋系統的示意圖，這個系統可彈性地配置以辨識出已經分析為 11 條影像分段的郵遞區號碼影像；

圖 10 A 和 10 B 共同顯示一個高階的流程圖，其顯示在根據本發明中第一個實施例之字符串詮釋方法施行期間被實施的步驟；

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(13)

圖 1 1 A 和 1 1 B 共同顯示一個高階的流程圖，其顯示在根據本發明中第 2 個實施例之字符串詮釋方法施行期間被實施的步驟；

圖 1 2 為本發明中之手攜式字符串詮釋系統之一示意圖；以及

圖 1 3 A 和 1 3 B 共同顯示一個高階的流程圖，其顯示在本發明中的字符串詮釋系統之訓練方法實行期間被實施之步驟。

本發明實施例的詳述：

如圖 1 所示，第 1 個實施例中的符號序列詮釋（亦即辨識）系統 1 包括若干個整合起來的組件。特別是，這個系統包含一個或多個處理單元 2（亦即微處理機），這些處理單元是由存放在程式記憶體儲存裝置 3 中的程式所控制。程式記憶體儲存裝置 3 亦存放著操作系統程式，應用程式以及各種本發明使用的影像處理副程式。資料儲存記憶體 4 被用來存放與本發明資料結構相關的資料。

一般而言，本系統具有一個輸入資料組擷取裝置 5，用以接收代表符號表象序列的一組輸入資料。在本實施例中，這種裝置，如圖 2 所示，可為接收灰階或彩色影像的影像檢測器，這些灰階或彩色影像可能是相連的阿拉伯數字字串，這些字串是記錄在一個記錄媒體 6 中。在此實施例中，字符串可為記錄在電氣被動式記錄平面，例如紙張，塑膠，毛品，織品等，或者記錄在一種電氣主動式記錄

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明(14)

平面例如業界熟知的壓敏數位化表面或「觸摸螢幕」LCD書寫及顯示表面。該字符串可以傳統的印刷或手寫(草書)方式被表達。

如圖1所示,本實施例中的系統具有隨機存取資料儲存記憶體(亦即VRAM),其用以暫存接收到將被詮釋處理的字符串影像。量大的資料儲存記憶體8被提供作為這些影像長久儲存之用。此系統最好亦具備一個裝有顯示幕或表面(LCD)的視覺顯示單元9;一個鍵盤或其他資料輸入裝置10;一種用來指向,拖曳及選擇被視覺地顯示在顯示幕上之圖形選項的裝置11;輸入/輸出裝置12;以及一個用以與一個或多個外部主系統14介面的系統介面13,其中的主系統14將使用到系統1提供的資訊。這些系統組件2,3,4,7及8最好能組裝在一個可攜式應用產品的外殼中。其他的組件亦具有它們各自的外殼。如圖示,這些組件以一種業界所知悉的方式藉由一個或多個系統匯流排15來與處理器2共同運作。在郵遞區號碼的辨識應用中,本發明的系統適用於與傳統的郵件分送機具14相介面,該機具14在郵政業務中亦屬已知。

如圖2所示,字符串詮釋系統1將執行多個功能,以便針對以圖形記錄下來的「字符串」完成詮釋處理,這項詮釋處理係基礎於取得之圖形化記錄字符串影像I中的畫素資訊分析,上述的「字符串」標識為Ci。影像處理的這些階段將依序在下文中參照其它指定的圖式被詳細說明

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明 (15)

大體而言，本發明提出的系統和方法可應用在任意長度的機具印刷或手寫的字符串。因此本發明在手寫辨識應用中將會很有用，其中使用者可在一種或他種書寫面上寫下一個或多個字，或者一個或多個文句，以供系統進行自動辨識。為清楚說明起見，圖 10 A 至 10 B 中所示之本發明的第 1 和第 2 實施例將考慮手寫郵遞區號碼的詮釋（分類）問題，其中，這些字符串的長度一般而言為 5 至 9 個數字。然而在其它實施例中，本發明所提出的方法和系統適用於詮釋任意長度的字符串（亦即文字），如同本領域中為大家所熟知的自動文句辨識系統中所提及的主題。

在圖 2 中，方塊 A 至 I 示意地代表了在本發明中的字符串詮釋處理其間被實施的各種處理步驟。如圖 2 中的方塊 A 所示，此處理之第 1 個步驟涉及擷取一個字符串的影像 I。基本上，被系統 1 取得的各個影像 I 包含像素組成的一個陣列。在此影像陣列中的各個像素具有一個灰階亮度值，此數值代表在此影像中該像素位置的影像強度值。此外，此像素的飽合值亦可被表示。各個取得的影像被存放在畫面（frame）緩衝器 7 中。如方塊 B 所示，該處理的第 2 個步驟涉及被儲存影像 I 的「前置處理」。由處理器 2 所執行的適當影像前置處理操作基本上包括：找出「欲處理的區域」的位置；去除底線；去除影像中歪斜及彎曲的部份；去除斑點（亦即相連的小部份）及加入必要的筆劃；並將影像正規化為一標準尺寸（亦即 20 個像素的

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明(16)

高度，而在寬度的選擇上則使得影像的橫寬比例不致改變)。應注意的是，影像正規化被完成使得經前置處理的影像 I' 可被傳送至系統的後續處理，而無需再作影像正規化。通常的情況下，在正規化處理期間執行的再取樣動作會產生一個實際上具有灰階的影像，即使原有的影像是黑白的。正規化影像的最上方和最下方輪廓然後被用來切除在影像中水平及垂直方向上字符的過長尾部。其他有關於上述影像前置處理操作的詳細說明可見 1991 年 12 月 31 日申請之標題為「阿拉伯數字影像分段機制」專利申請序號第 07 / 816, 414 號。

方塊 C 所示辨識程序的次一步驟係將前置處理後的影像 I' 切割為多個子影像，稱為「細元」。產生影像細元的目的是使得影像細元能夠結合起來，如此在方塊 D 所示之影像分段格式化步驟進行期間，可以構成所謂的影像「分段」S_i。

根據本發明，影像細元是被首先對前置處理完之影像執行「連結成份分析」以檢知出大型「連結成份」之存在所產生的。之後，針對包含大型連結成份的子影像，使用一種「智慧型」擺動切割線繪圖程序。連結成份分析和智慧型擺動切割線繪圖兩種子程序是藉由程式化的處理器 2 搭配相關 RAM 4 加以完成。更確切而言，連結成份分析針對包含了前置處理影像的像素強度進行分析，以便確定連接在一起的字符成份（亦即像素群）之存在。應注意到連結成份分析係沿著影像的水平及垂直方向搜尋具有預設

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

康

五、發明說明(17)

強度值範圍的像素群，以及能構成相連在一起的大型字符成份，以及找出那一個像素群可能與一個或多個字符相關連。圖4 A至4 F中所示之第2及第3郵遞區號碼影像舉例說明了連結字符成份。

對於一個具有大型連結成份的子影像而言，有可能包含一個以上的字符。重要的是能確定出在該子影像中應繪出的切割線，而使得僅有一個字符可由一個影像細元的畫素來表示。這是藉著在指定之大型連結成份中產生「擺動」切割線的方式來完成上述目的。一般而言，此切割線產生程序可僅需在代表字符之像素群中繪出切割線，而將由大型連結成份代表的字符劃分為二個或多個影像細元。可將相鄰影像細元結合起來以構成一個影像分段的方法數目將隨著在這個辨識步驟中所產生之影像細元數目之增加而大幅增加。

本發明中的系統避免將前置處理後之影像切割為過小的影像細元，方法是利用複雜的啟發式訓練，以指定出一組好的切割線，去除多餘的線條或物件。圖4 A至4 F中，針對該子程序操作的說明，在前置處理後的影像中繪出切割線以及選擇性地移去切割線。在子程序結束時，剩餘的各對相鄰切割線間的像素定義為一個影像「細元」，圖5列出所產生的影像細元。這些細元是在影像細元範例產生期間所求得的，如此表所示，各影像細元由一個細元數字，亦即0，1，2，3，4等予以標識，有關於在辨識程序執行期間自動產生切割線的步驟可詳見申請專利序號

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(18)

N o . 0 7 / 8 1 6 , 4 1 4 的內容。

如圖 2 中的方塊 D 所示，後續的處理步驟係將連續的影像細元依由左向右的次序結合起來以求得一組影像“分段”，如圖 6 中的列表所示。藉由將指定給各影像分段構成影像細元的數字，例如 0，0 1，1，2，2 3 等結合起來，可標識出各個影像分段。理想狀況下，各個影像分段包含代表一個且僅有一個字符的像素。然而，事實並非總是如此。重要的是，最後一組影像分段應包含正確的影像分段。複雜的啓發式訓練被用來確定有多少影像細元以及那些影像細元應結合起來以構成影像分段。一般而言，啓發式訓練是被描述為“限定”切割，外在互連成份“切割”，“內在互連成份”切割等。這些啓發式訓練的參數和調整因子是經由實驗而確定。各個影像分段包括一組影像像素，這一組影像像素將被系統援用的指定神經資訊處理網路予以分析。如同以下內容的詳述，各個神經網路的功能係用以分析各個影像分段的像素集合，其中前述的影像分段是由系統指定，且會產生分析之輸出結果，這個輸出結果即為各個（1 0）可能數字字符（即，符號）的得分，而上述的像素集合可能就是代表這個符號或被歸類為可能的集合。

本程序的下一個步驟概略地被示於方塊 E 中，其將連續的影像“分段”依由左至右之次序結合在一起以便得出一組可接受的（亦即合法的）影像“串接分段”。各個影像分段應考慮到前置處理影像 I 中的所有像素。所考慮

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

康

五、發明說明(19)

到的影像串接分段最好是儘可能地少，但仍確保正確的串接分段是包含在所有影像串接分段所成的集合中，而這些影像串接分段是由已產生的影像細元所構成。在圖 7 的列表中，針對 5 個字符郵遞區號碼說明之許多合法的影像串接分段例，顯示出其中 3 個。如方塊 E 所示，串接分段是使用圖 8 中「直接非循環對正圖形」隱涵性地構成。這種圖形結構確保各個影像串接分段能夠包括 5 個影像分段。爲了擷取輸入影像 I 的空間結構實體，將存在若干法則，作爲影像分段可串接在一起的依循。例如，一個分段的右緣必須位於後接影像分段的左緣附近（亦即不允許跳過多個像素以及／或者以錯誤的空間次序將像素結合起來）。然而，若有需要，若干此類限定在某些狀況下可從寬考慮。有關串接連續影像分段 S_i 的其他詳情被揭示於上文所述與本案，同時係屬中之專利申請序號 No. 07 / 816,415 標題爲「影像辨識系統中用以自動分段及辨識的繪圖系統」。若有需要，選擇出來的影像之串接分段可被明確地展示在方塊 F 中。

應注意的是，本發明中的直接非循環圖形亦提供一種新式的裝置，能夠同時針對前置處理影像 I 的可能影像串接分段 { S } 集合以及字符串詮釋集合 { C }（亦即分類）集合兩者進行模型化（亦即表示），其中，這些字符串詮釋 { C } 是由字母所許可且可能利用語言或碼的詞語加以限定，其中，記錄的字符串已被表象出來。如同配合圖 8 所作的詳述，本發明中的系統使用一種以圖形方式表

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

康

五、發明說明(20)

示為「直接非循環圖形」的資料結構，以整合的方式，將影像串接分段和字符串詮釋問題化約為「圖形中最佳路徑」的問題。直覺上，此種問題的化約具有幾何的訴求。以下將詳細說明排列圖形，實現此圖形的資料結構，以及使用它的程序。之後，將特別詳細說明使用此圖形的程序，如在圖2中方塊G所示之影像分段分析階段期間，方塊H中所示路徑得分及機率計算階段，以及方塊I中所示字符串詮釋階段。

如圖8所示，本發明中的圖形包含一個二次元的節點陣列，以一個高階描述的觀點來看，這個節點陣列類似習知技術中的稱為格子或格架圖的圖形。如同以下的詳述，本發明中的排列圖形是用一個資料結構來實現，這個資料結構執行多個重要的模型化功能。此資料結構是以一種在程式撰寫技術中為已知方法被程式化的處理器2所產生，修改及管理的。在排列圖形中的各個節點可實現為一個個別的資料結構，這是一個「主資料結構」的子結構。各節點的資料結構具有多個「區域」資訊欄位，專門用以標示及儲存以下的資訊項目：一個單一的節點辨識碼（亦即用以標識該節點行／列位址的碼）；可由與影像分段相關的像素所代表之各個可能數字符號的計算得分；可由影像分段相關的像素來表示之各個可能數字符號的「非正規化」計算得分；先驅節點的節點辨識碼；以及其後繼節點的節點辨識碼。

為了儲存在各處理階段中所產生的資訊，主要資料結

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明 (21)

構具有多個「總體」資訊欄位，專門用以標示及儲存以下的資訊項目，包括：一組用以辨明那一個特定影像分段是由排列圖形中各特定列之節點所代表的辨識碼；一組用以辨識出各個影像分段存放於記憶體位置的位址；以及沿著之排列圖形中被選取的路徑和代表相同的字符串詮釋通過該排列圖形之路徑組群的得分總和。

在第 1 說明實施例中，排列圖形中的行 (columns) 數等於在可能字符串詮釋中的字符數 (亦即圖 3 中的 5 個郵遞區號碼)。此外，排列圖形中的列 (rows) 數等於在影像分段產生處理階段中建構出的影像分段數。因此，排列圖形的大小基本上隨著各影像 I 中列長度之不同而有所變化對字符串詮釋 (亦即分析和分類) 而獲得。因此，對於各個取得的影像 I 而言，程式化處理器 2 週期性地產生圖 8 中所示型式的圖形，這種圖形特別適用於取得之影像。每一個這樣的排列圖形是藉著產生存放在 RAM 4 中的資料結構而被實質上地實現。與影像 I 之影像分段相關的資訊以及其所有的可能字符串詮釋係被存放在特別為此影像所產生的資料結構中的資訊欄位中。最後，程式化處理器 2 使用組合起來的資訊以便自候選的集合 {C} 中選取最合適的字符串詮釋 C。

如圖 8 所示，本發明中的排列圖形具有多個不錯的結構化特點。圖形中的主要部份包含列和行。各行對應於字符串詮釋 C 中的一個字符位置。因為這個例子係處理 5 個字符的郵遞區號碼，因此如圖所示需要 5 行。各列對應於

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(22)

一個影像分段。因為這個例子包含 11 個分段，如所示需要 11 列。每一列和每一行的交點即成為一個節點，以一對黑點（· ·）來表示。左邊的那一點代表節點的「晨」部，而右邊這一點則代表節點的「夜」部。各個節點可以其列指標和其行指標來指定。此外，存在一個特殊的起始節點 17，位於第 1 字符位置之前，以及最左影像分段的左邊。亦存在一個特殊的末端節點 18，位於最終字符位置的右邊以及在最右影像分段之下方。

如圖 8 所示，存在 10 個辨識弧線，連接著各節點的晨部和夜部。為清楚說明起見，這 10 個辨識弧線中僅有 3 個見於圖 8 中。在詮釋處理期間，各個辨識弧線 19 標示著一個「r-得分」，它被指定給由辨識弧線表示的字符。在此實施例中，這些辨識弧線代表被指定給包含遞區號碼的數字字符之未經正規化的 r-得分。然而，在文字和文句的辨識應用中，這些辨識弧線基本上代表著被指定給一個預定字母或字彙中的符號之未經正規化的得分。如圖 8 所示，一定向的膠連弧線（glue-arc）19 亦出現在某一節點의各個夜部和其緊鄰節點의晨部之間，以代表這些節點中的前系節點和後系節點。與辨識弧線不同的，這些膠連弧線在本實例中並未經由神經網路指定 r-得分。在其他的實施例中，可使用複雜的膠連弧線得分，但是對本實施例而言，可使用較簡單的系統：被允許指定給弧線的得分為 1.0 並維持，然而不允許的弧線被指定得分 0.0，並將其自排列圖形中剔除。

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明(23)

如圖所示，任何節點的晨部可具有一個以上的膠連弧線進入此節點。同理，任何節點的夜部可具有一個以上的膠連弧線離開此節點。由於對影像串接分段的構造有所限定，因此在排列圖形中，存在著區域性有意義的膠連弧線，但是對總體性而言沒有意義，因此，可去除或刪減若干膠連弧線以便改善詮釋處理的計算效率。在進行圖2中方塊G所示之影像分段分析階段之前，可針對排列圖形執行以下的「膠連弧線」刪減程序。

膠連弧線刪除程序的第一個步驟是計算節點的「前向錐」，這些節點是起始節點の後繼節點，計算的方法是互動地（interactively）針對已標示為前向錐之成員的節點之後繼者進行標示。第2個步驟計算出節點的「反向錐」，這些節點是末端節點的先驅節點，計算的方法是針對已標示為反向錐之成員的節點後繼者進行標示。第3個步驟是決定那些節點不在這二個錐體的邏輯交點上，然後將這些錐體標示為「死亡」。之後，任何延伸至「死亡」節點或來自「死亡」節點的膠連弧線均自許可膠連弧線的列表中剔除（刪掉）。存在於這些錐體交點的各個節點被視為「存活」，且在影像分段分析階段，將會有得分被指定給其辨識弧線組。應注意的是，藉著滿足總體限制，基本上在排列圖形中的右上及左下角落處存在有許多節點是不具有合法的先驅節點或後繼節點。如圖8所示，上述事實可由排列圖形的這些區域中，未存在進入及離開的膠連弧線所表示。此外，若有需要，可使用限定切割的存在來刪

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

線

五、發明說明(24)

減排列圖形。

圖形中的各條路徑代表了一個串接分段和一個詮釋兩者。在路徑上的膠連弧線指定一個串接分段，而路徑上的辨識弧線則指定一個詮釋。爲了了解本發明之程序如何自整個可能字符串詮釋集合，或者自一個很小的競爭字符串詮釋集合中，選取“正確”的字符串詮釋，有必要先了解幾個最後用以選取“正確”字符串詮釋的子程序。第1個子程序係關於指定給各節點處的辨識弧線之非正規化 r -得分的計算。第2個子程序係關於與所有代表了相同字符串詮釋之排列圖形中的字符串路徑有關之 r -得分的總和之計算。以下將說明這些子程序。

如圖9所示，詮釋處理中的影像分段分析階段將使用一種神經計算網路21的複合體。各個第 i 神經計算網路的主要功能是用以分析影像分段 S_i 的像素，而此影像分段 S_i 是與圖形中的第 i 列共同加以索引，並計算一組指定給位於圖形中第 i 列的各個節點之辨識弧線的“得分”。

對每一列而言，存在一個分段以及一個神經網路，且在同樣一列中的所有節點接收了一組相同的10個 r -得分。爲說明起見，對於各個節點而言，圖9中僅顯示10個辨識弧線中的3個。重要的是，各個神經計算網路將其輸入（一群由一組數字代表的像素）映對至一組稱爲 r -得分的10個數字 $r_0, r_1, \dots, r_9$ 。網路中的架構保證這些 r -得分爲正值，容許它們的詮釋爲非正規化的機

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

線

五、發明說明(25)

率值。一個大的 r 數值代表輸入分段代表著數字 '0' 的一高機率值，相同地，其他的 9 個 r - 得分分別對應於其他的 9 個數字。一個大的 r - 得分值亦反映出輸入分段是影像中一個正確串接分段的一部分的一個高機率值；相反地，若一個分段是藉著將一數字切為一半以構成的話（有時會發生這種狀況），該分段的所有 10 個 r - 得分值將會很小，如此可代表偵測到分段的不正常特性。

根據傳統，各個神經計算網路的映對功能是以一組可調的參數來特徵化，這些參數可以向量形式來表示，亦即具有分量 $w_1, w_2, \dots, w_m$ 的加權向量 w 。一開始，該組可調參數的各神經計算網路被調整為一組初始值。然而，如同以下的詳述內容，將提供一個圖 2 中方塊 J 所顯示的神經網路參數調整階段，如此在一個或多個學習程序中，這些參數可遞增性地予以調整，進而使得各個神經計算網路的輸入／輸出映對功能可符合一組可靠的訓練資料。在本發明的較佳實施例中，這個訓練資料集合包含幾十萬個代表郵遞區號碼的有效訓練影像，而這些郵遞區號碼是由國內不同人所手寫出來的。

在此實施例中，由各個第 i 神經計算網路所產生的 r - 得分可以向量型式 $r = r_1, r_2, \dots, r_N$ 來表示，且被指定給位於排列圖形的第 i 列中的所有節點處的 10 個對應的辨識弧線（亦即資訊欄位）。一般而言，各個神經計算網路可實現為一個計算機程式，一個電子電路或某些能夠實現成神經計算網路的輸入／輸出映對功能的微型及

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(26)

巨型元件。然而，在此較佳實施例中，各神經計算網路是藉由執行爲大家熟知的 LeNet™ 計算機程式而予以實現，詳情可見 Y. Le Cun 等人所著之「以反向傳遞網路進行手寫數字辨識」一文，此文見於 1990 年 (Morgan Kanfman) 由 David Touretzky 編著的「Advances in Neural Information Processing 2」，396-404 頁。有關於神經計算網路訓練及結構的進一步討論可見「自動化學習，法則抽取以及歸納」一文，由 John Denker 等人所著，出版於「Complex Systems」書中第 877-922 頁，vol. 1，1987。

在排列圖形中，可以有二條或多條代表相同字符串詮釋的路徑（代表不同的串接分段）。根據本發明的原理，代表一特定詮釋的路徑將視爲一「群 (group)」。指定給一特定詮釋的得分將與群中所有路徑的得分總和有關。此狀況與習知的辨識器不同，習知的辨識器大體上僅考慮群中一條路徑的得分，而不考慮群中其他路徑的影響。

對於具有 5 個數字的影像而言，存在大約有 10^5 個可能的不同詮釋，而排列圖形中的路徑數目可能更大。明確地顯示這些路徑及詮釋，或者個別考慮各機率值並不實際。本發明中的資料結構和演算法可使系統辨識出若干重要的路徑群，例如對應於一特定詮釋的路徑群，或者包含所有路徑的路徑群，同時有效地評估該群的得分，亦即在群中路徑得分的總和。

本發明中的系統分析取得影像 I 的像素，並且計算圖

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(27)

形中所有路徑的總和，其中的圖形代表候選詮釋（亦即分類），據此以計算出機率值。總和中的各項是指定給弧線之得分的乘積，而這些弧線構成排列圖形中的某一路徑。正規化僅在總和計算得出後才執行。吾人稱此為「字串接著字串」的正規化。相反地，計算機率值的習知辨識器大體而言是在處理期間極早的時間點進行得分正規化，基本上與「逐字」正規化的方式相同，因此丟棄了串接分段品質相關的有用資訊。重要的是，以下提及的神經計算網路訓練程序訓練神經網路的複合體以產生 r - 得分，這個 r - 得分含有關係到一串接分段是正確的機率值資訊，而不只是一個分段之特定字符詮釋是正確的機率值。

利用本發明中之系統和方法所產生的正規化得分代表著由果溯因機率 $P(C/I)$ 的預估值。相反地，習知技術中多字符辨識 MCR 系統所使用的最大可能序列預估機率一般使用的是 $P(I/C)$ 型式的前視機率（亦即可能性）。對於許多的目的而言，這是可以接受的，因為這些不同的機率量度在給予某些額外的資訊下可彼此相關。由果溯因格式的真正的優點為本發明之系統和方法的內部計算是與詮釋及串接分段的由果溯因聯合機率預估值 $P(C, S/I)$ 有關。對應的前視（可能性）表示式 $P(I/C, S)$ 無法容易地與有用的由果溯因格式取得相關，因為廣泛地預估邊際機率 $P(S)$ 並非永遠可行。因此，習知的辨識器基本上不能指定未適當予以正規化的得分，雖然它們能辨識最高得分詮釋。本發明中經適當正規化的得

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明 (28)

分可較輕易地詮釋為機率，進而可以較輕易地與取自其他來源的資訊結合起來。

大體而言，圖 10 所示之程序的目的是針對各個競爭字符串詮釋，計算出一個新的由果溯因機率值 $P(C/I)$ ，其中的競爭字符串詮釋是由圖 9 中的排列圖形所表示。應注意的是，各個機率的計算值是一個具有分子部份和分母部份的比值。以數學的觀點而言，本發明中的機率量度可表示為：

$$P(C, I) = \frac{N(C/I)}{D(I)} = \frac{\sum_{S' \cdot I} \Pi r(C_i, S'_i, I)}{\sum_{C' \cdot S'' \cdot I} \sum \Pi r(C'_i, S''_i, I)}$$

應注意的是，分子部份的第 1 項 $\Pi r(C_i, S'_i, I)$ 表示沿各路徑 (S'_i) 之弧線中 r - 得分的相乘級數，

而整個分子部份 $\sum_{S' \cdot I} \Pi r(C_i, S'_i, I)$ 代表所有路徑（亦即串接分段 S' ）中該種路徑得分乘積的累加值，上述所謂的「所有路徑」係代表相同的字符串詮釋。在分母中的

第 1 個表示式 $\sum_{C' \cdot S'' \cdot I} \sum \Pi r(C'_i, S''_i, I)$ 表示所有路徑中路徑得分乘積之總和，而其中所謂的「所有路徑」代表相同的字符串詮釋，而整個分子部份 $\sum_{S' \cdot I} \Pi r(C'_i, S''_i, I)$

代表所有字符串詮釋 $\{C\}$ 之所有路徑得分乘積的一個總

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明 (29)

和，其中的字符串詮釋 { C } 表示排列圖形。

因為分母部份包括所有可能詮釋的貢獻值，這個數值僅與取得的影像 I 有關，而與特定的詮釋 C 無關。分母部份的目的是確保機率值能適當地正規化，如此，依據機率的一般性原理， $P(C_i / I)$ 的總和（對所有 C_i 而言）等於 1。

一旦針對一個特定的字符串詮釋計算出分子部份，只需將共同之分母來除計算得出之分子，則可取得該字符串詮釋的機率。幸運的是，存在著多個不同的方法可供上述機率計算程序使用，例如可使用較嚴謹的程序，以獲得一個「正確」的字符串詮釋。圖 10 A 和 10 B 中的流程圖顯示一個方法，而在圖 11 A 和 11 B 中的流程圖則說明了另一個方法。以下將針對這二個方法作說明。

圖 10 A 和 10 B 的流程圖顯示出本發明中第 1 字符串詮釋程序的步驟。如方塊 A 所示，本程序的第 1 個步驟係使用一個圖 9 中所示之第 i 神經計算網路來針對圖形中第 i 列的各個節點，計算出一組 r - 得分。然後，如方塊 B 所示，這個程序使用為大家熟知的 Viterbi 演算法來辨識（如為代表膠連弧線及辨識弧線的一串列碼般）出排列圖形中具有最大的路徑得分的路徑。之後，處理器辨識出對應於此路徑的字符串詮釋。因為這個字符串的路徑得分僅是一種近似，本身並非一種可靠的量度，因此予以捨棄。僅有由該路徑（例如 5 個字符的郵遞區號碼 3 5 7 3 3）所表示之辨識字符串詮釋 C (v) 的資訊可被保留下來

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(30)

接著，如圖 10 A 中的方塊 C 所示，本程序使用著名的「前向演算法」來針對辨識出的字符串詮釋，計算出機率量度的共同分母部份 $D(I)$ 。這個數值然後被存放在一個被用來實現圖 8 中的排列圖形之主要的資料結構中。前向演算法的使用係針對所有以排列圖形表示的可能字符串詮釋，提供非正規化 r -得分的乘積（沿各條路徑）總和（對所有路徑而言）準確值。在某些實施例中，膠連弧線僅藉由是否存在而決定對路徑得分的貢獻。在某些更複雜的實施例中，得分可被指定給膠連弧線（亦指定給辨識弧線），同時這些所有的得分均包含在各條路徑中乘積的係數中。

如圖 10 A 中的方塊 D 所示，在計算出上述的共同分母部份 $D(I)$ 之後，本程序使用「前向演算法」來針對先前以 Viterbi 演算法辨識出的正確字符串詮釋 $C(v)$ ，計算出機率量度的分子部份 $N(C(v), I)$ 。這個數字再存放到用來實現圖 8 中的排列圖形之主要資料結構中。在此，前向演算法接收由 Viterbi 演算法所辨認出的選取字符串詮釋之數碼作為輸入，且針對這個選取的字符串詮釋 $C(v)$ ，求出一個準確的分子數值（亦即一個有限的總和）作為輸出。應注意的是，針對字符串詮釋所計算求得的分子部份等於沿排列圖形中各路徑之非正規化 r -得分的乘積總和（對所有路徑而言），而上述的排列圖形係代表了字符串詮釋 $C(v)$ 。在分子部份的計算期間，

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明(31)

膠連弧線的處理與分母計算期間的處理是相同的過程。

如圖 10 A 中的方塊 E 所示，當分母部份和分子部份計算求得後，可針對字符串詮釋 $C(v)$ 計算求得改善的機率值 $P(Cv/I)$ 。這個機率值則被存放在主要資料結構中。最後，如圖 10 B 的方塊 F 所示，處理器確定在方塊 E 中計算求得的機率值是否大於一個限值。若是，則處理器確知由 Viterbi 演算法選出的字符串詮釋對於被分析的影像 I 而言是具最高機率值的字符串詮釋。接著在方塊 G 中，處理器由此系統中產生 (i) 字符串詮釋（亦即 3 5 7 3 3）以及 (ii) 計算求得的相關機率兩者作為輸出。結合此兩個輸出值可作為決定如何發送此郵件的依據。

在處理的這個階段中，可能存在著一個或多個理由來說明對以這種方式執行額外運算以辨識更多高得分詮釋是有裨益的。其一是吾人希望即使方塊 F 中指定給 $C(v)$ 的機率值是小於 0.5，也能確保可辨識出具最高機率值的詮釋。在此狀況下，一組競爭字符串詮釋被辨識出，同時針對這組詮釋中的各字符計算出其機率值。此外，本發明亦可作為某一大型系統的一部份，在此大型系統中，多個詮釋（以及機率值）可應用在後續的處理中。特別是，一個在本發明中係依據取到的畫素影像被給予高機率值的詮釋可能在此大型系統中的後續處理階段中被排除，針對以上狀況，有必要存在另外一個詮釋。

針對這些或其他理由，可應用圖 11 A 及 11 B 中流

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

東

五、發明說明 (32)

程圖所說明的替代程序。

如圖 1 1 A 中的方塊 A 所示，此程序中的第 1 個步驟步涉及使用第 i 神經計算網路以針對圖形中第 i 列的各節點，計算出一組 r - 得分。之後，如方塊 B 所示，此程序使用一種射線搜尋演算法來辨識（作為代表膠連弧線及辨識弧線的一串數碼）出排列圖形中極小的一組路徑。之後，此組競爭字符串詮釋 $\{C_j\}$ ，及對應於此組的路徑被辨識出。

如圖 1 1 A 中的方塊 C 所示，處理器使用了著名的前向演算法來計算出分母 $D(I)$ ，這個 $D(I)$ 是作為在該組競爭詮釋 $\{C_j\}$ 中的各個詮釋 C_j 之機率值 $P(C_j, I)$ 的分母部份。這個數值係存放在主要的資料結構中。再者，前向演算法針對未正規化 r - 得分之乘積總和（對所有路徑而言），提供一個準確值。而上述的非正規 r - 得分係針對沿各路徑的弧線所求得，在此分母部份而言，總和值的計算係考慮到所有可能的詮釋。

為了計算辨識出詮釋的得分，如方塊 D 所示，處理器使用前向演算法來計算各個競爭字符串詮釋 C_j 之機率值的分子部份 $N(C_j / I)$ 。這些數字則被存放在主要資料結構中。這個前向演算法針對未正規化 r - 得分的乘積總和（對所有路徑而言），提供一個準確值，其中的未正規化 r - 得分係針對各路徑的弧線所求得。

應注意到利用前向演算法計算求得的總和是考慮到路徑的一個總和值。所謂的一條路徑是指在方塊 B 中由射線

（請先閱讀背面之注意事項再填寫本頁）

訂

原

五、發明說明(33)

搜尋演算法辨識出的路徑。事實上，這個路徑產生總和值中的最大項。若某人可假設該總和值可由其最大項獲得不錯的近似，則沒有必要執行前向演算法以導出分子；射線搜尋演算法取得的結果已足夠使用。此狀況稱為「單項總和」近似。然而，總和值並非總是可以其最大項作為不錯的近似，而比較可取的作法是排除由射線搜尋演算法計算出的得分，但是保留由射線搜尋演算法辨識出的詮釋，同時利用前向演算法求出該保留詮釋的得分。

基本上沒有必要針對所有的可能詮釋計算出分子部份，這就是為什麼最好能夠辨識出一組極小的詮釋（如方塊B），而因為存在有大的「單項」得分，上述的詮釋應具有大的分子數值，而由此可獲得大的機率值。

以上的討論說明了系統被訓練之後的操作情形。現在將說明系統的學習模式。

為了獲得最佳的性能，本發明中的字符串詮釋系統具有一種操作的自動化學習模式，其可使系統在一個或多個學習階段，被自動地被訓練。這種操作模式將配合圖2中的方塊J以及圖9中的系統加以詳述。

如圖2中的方塊J和圖9所示，本發明中的字符串詮釋系統包括一個神經網路參數調整模組29，這個模組29將與圖9所示系統之圖形30和神經計算網路21的複合體起交互作用。一般而言，本發明中的訓練程序係根據監督式學習的概念，換言之，對訓練集合中的各個影像I*而言，存在一個詮釋C*與之對應。在此說明實施例中

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

東

五、發明說明 (34)

，神經網路參數調整模組的設計係確保正確字符串詮釋的期望（亦即平均）機率值 $P(C^*/I^*)$ 在訓練集中所有影像 I^* 的處理期間能夠增加，而各個不正確之字符串詮釋的期望機率值 $P(C/I)$ 在訓練過程中能降低。簡言之，學習模式以及參數調整模組的作用是確保錯誤詮釋的平均機率值為最小，而正確字符串詮釋 C 的機率值為最大。在本說明實施例中，選取 $\log[P(C/I)]$ 作為目的函數，因為 $\log$ 函數在接近零的範圍較為陡峭。如此使得訓練程序能強調低得分像素樣式（亦即影像分段），這是較為有利的方式，因為這些低得分像素樣式在處理上較為棘手，且需要最多的訓練。

為了實現這個選出的目的函數，處理器使用梯度，可表示為：

$$\frac{\partial \log P(C/I)}{\partial w} = \frac{\partial \log P(C/I)}{\partial r} \cdot \frac{\partial r}{\partial w}$$

其中 $w = w_1, w_2, \dots, w_m$ 為 m 次元神經網路加權向量，而 $r_i = r_1, r_2, \dots, r_n$ 為第 i 個 n 次元 r - 得分向量，而此 r - 得分向量是第 i 個神經網路所產生的輸出。一般而言，加權向量 w 具有 10, 000 或更多的分量。這個 r - 得分向量 r 正好具有 10 個作數字辨識的分量。對上述梯度表示式右邊所作的點積（dot product）代表著 r 分量的總和。一般而言，對各個神經網路，亦即對於排列格子圖中的各列而言，存在一個這樣

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

東

五、發明說明(35)

的梯度表示式。有時候，使用相同的加權向量 w 來控制一個以上的網路是較為有利，在這種情形中， w 的梯度包括來自上述各網路的貢獻。在本實施例中，如圖 9 所示，加權向量是存放在一個暫存器 31 中，這個暫存器

31 將相同的加權向量提供給系統中的每一個神經網路。

在啟動此處提及的多字符辨識器訓練處理之前，必先賦予神經網路加權向量初始值。這可使用根據一些合理分配的隨機值來進行初始化，或者使用選出的前視數值來初始化，如果這個值特別適用。在許多狀況下，最好是暫時性地切斷神經網路與排列圖形的連接，且針對手寫分段影像，作事前的訓練，如同被使用成一個單一字符辨識器一般。得出的加權向量值作為上述多字符辨識器訓練處理的起始點。

梯度表示式的左邊稱為系統靈敏度向量，因為這個梯度提供有關於整個系統對應於加權向量 w 的改變之輸出靈敏度的訊息。系統靈敏度向量的各個分量屬於加權向量的對應分量。特別是，如果系統靈敏度向量的一特定分量大於零，則加權向量中對應分量的小幅增加量將使得機率值 $P(C|I)$ 有所增加，而此機率值 $P(C|I)$ 是由系統針對問題中之影像 I 而指定給詮釋 C 。簡言之，系統靈敏度向量可用以針對上述的目的函數作最佳化處理。

為了對訓練處理的原理有一更深入的理解，針對組成梯度函數的數量特性深入了解是有助益的。

根據以上的公式，系統靈敏度向量經過計算後成為公

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(36)

式右邊顯示之其他二數量的點積(向量矩陣積)。其中的

第一個數量是一個向量 $\partial \log P / \partial r$ ，它提供了圖形輸出靈敏度相關的訊息，而其中的靈敏度對應於 r - 得分輸出 $r_1 \dots \dots r_n$ 的變化量。它可視為圖形靈敏度向量。第 2

個數量是一個 $N \times N$ 矩陣 $\partial r / \partial w$ ，它提供了神經網路輸出靈敏度相關的訊息，而此靈敏度係對應於控制所有神經網路之加權向量中的變化量。

若僅作概念上的說明，上述三項彼此在函數中的關聯性可見以下的內容。在各個訓練影像 I^* 的處理期間，所求得的神經網路靈敏度矩陣用以轉換求得的圖形靈敏度向量，如此可產生一個評估後的系統靈敏度向量。之後，經評估的系統靈敏度向量調整加權向量中的對應分量，如此參數調整模組中的目的函數 $P(C^* / I^*)$ 可得到最佳化。

理論上，可利用數值方法求得梯度函數右邊各項的值，之後執行該處指定的數學運算。然而，在訓練階段期間，存在一種較簡單的方法，能輕易地求出各個影像 / 詮釋對 $\{I^*, C^*\}$ 的系統靈敏度向量。如以下內容之說明，並且參考圖 13 中的流程圖，可使用著名的後向傳遞(Back-Prop)演算法配合有效的計算方式以求出系統的靈敏度向量，而無須明確地求出神經網路靈敏度矩陣。

當本發明中的系統是工作在學習模式時，如圖 2 中的

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

東

五、發明說明(37)

方塊 K 所示，可針對訓練集合資料庫中的各個影像 I^* 來執行圖 13 中的訓練程序。各影像 I^* 均與一個已知的字符串詮釋 C^* 相關聯。基本上，在特定訓練階段期間，可使用大量的（幾萬個）影像／詮釋對 $\{I^*, C^*\}$ 來訓練系統。如圖 2 中的方塊 B 所示，各影像 I^* 基本上使用與本發明中詮釋處理期間相同的方法來進行前置處理。同時，如圖 2 中方塊 C 至 E 所示，基本上使用與本發明中詮釋處理期間所使用的相同方法來分別針對影像 I^* 建構影像分段和影像串接分段。

之後，如圖 2 中的方塊 F 所示，將針對產生出來的影像串接分段和與影像 I^* 相關的可能字符串詮釋，建構一個圖形模型。在訓練程序的這個階段裡，本說明實施例中的訓練方法顯現出以下的事實。首先，各個機率 $P(C^*/I^*)$ 具有一個分子部份 $N(C^*/I^*)$ 以及一個共有的分母部份 D_I 。其次，利用已知的微分和對數特性，對應於 r -得分變數的圖形靈敏度向量（亦即 $\partial \log [P(C^*/I^*)]$ 的偏微分）可表示為：

$$\frac{\partial \log P(C/I)}{\partial r} = \frac{\partial \log N(C/I)}{\partial r} - \frac{\partial \log D(C/I)}{\partial r}$$

重要的是，在此等式左邊明顯表示出來的圖形靈敏度向量可輕易地以圖 13 A 及 13 B 中的程序予以求出，以下將有所說明。

如圖 13 A 中的方塊 A 所示，處理器執行前向傳遞演

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

象

五、發明說明(38)

算法一次，可以數值方法針對影像／詮釋 $\{I^*, C^*\}$ 求出機率值 $P(C^*/I^*)$ 的分子部份，並且執行一次以求出其中的分母部份。這些數值將被儲存起來。重要的是，在處理期間，前向演算法明確地顯現出針對影像／詮釋對 $\{I^*, C^*\}$ 所建構的圖形隱含地表示了分析得到的（亦即代數上的）表示式，這些表示式以數學形式表示出相關機率值 $P(C^*/I^*)$ 的分子部份和分母部份。

在圖 1 3 A 中的方塊 B 中，處理器執行習知的 Baum-Welch 演算法以便以數值方法求出機率值 $P(C^*/I^*)$ 中分子部份的偏微分，而這個偏微分係指對變數 r 作偏微分。在方塊 C 中，處理器利用前向演算法計算出機率值 $P(C^*/I^*)$ 的分母部份數值。在方塊 D 中，處理器執行習知的 Baum-Welch 演算法以便以數值方法求出機率值 $P(C^*/I^*)$ 中分母部份對變數的偏微分。之後在圖 1 3 A 中的方塊 E 中，處理器使用求得的分子和分母部份，以及它們的偏微分值，以便根據上述的公式以數值方法求得圖形靈敏度向量。

爲了有效地針對影像／詮釋對 $\{I^*, C^*\}$ 求得系統的靈敏度向量，訓練程序將設定各個神經網路的輸出層梯度向量，如圖 1 3 B 中的方塊 F 所示，等於以數值方法計算而得之圖形靈敏度向量中的對應分量。之後如方塊 G 所示，處理器使用後向傳遞演算法且根據上述的公式計算出系統靈敏度向量，利用後向傳遞處理以計算所需結果的細節可見文獻“自動學習”，法則抽取以及產生”由 Denker

(請先閱讀背面之注意事項再填寫本頁)

訂

原

五、發明說明(39)

等人所著。

應注意的是，後向傳遞演算法並非外顯地用來計算出神經網路靈敏度矩陣，而是用來求出神經網路靈敏度矩陣以及圖形靈敏度向量的向量-矩陣積。這個結果是整體系統靈敏度向量的外顯導出值。後者針對各個神經網路中加權向量的各個分量，提出一種更新用的有利方向。如方塊H所示，在各影像I*的處理過後，處理器使用系統靈敏度向量的各個分量以更新加權向量中的各個分量，而上述的系統靈敏度向量是以數值方法求得。以下將說明較佳的更新程序。

在更新進行之前，加權向量中的各個第i分量是以 w_i 表示，而在更新之後，各個第i分量是以 w_i' 表示。在各影像I*處理之後，加權向量則依據下列的表示式作更新：

$$w_i' = w_i + \delta_i \frac{\partial \log P(C^*/I^*)}{\partial w_i}$$

其中的 δ_i 是“步級尺寸控制參數”，其中 w_i' 則表示更新後的加權向量，而 $\frac{\partial \log P(C^*/I^*)}{\partial w_i}$ 則是 $\log P(I^*/C^*)$ 對 w_i 的偏微分。理論上，存在多個不同的步級大小控制參數 δ_i ，其中的每一個參數均與加權向量中的一個分量相關，但事實上，將所有參數 δ_i 值設為相同則較為方便。一般而言，步級大小控制參數值與以下二者有關(i)針對輸入至神經網路之像素所選用

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

原

五、發明說明(40)

的正規化因數，以及 (i, i) 針對神經網路中間數值所選用的正規化因素，（所謂的中間數值係指神經網路中由一層至下一層的輸出值），而這些步級大小控制參數可在訓練階段再作預估。

基本上，對於步級大小控制參數適當數值的選用，有二個主要的考慮。若此控制參數的選用值過小，則加權向量 w 收斂至一最佳值的速度過慢。另一方面，若選用的控制參數數值過大，則存在相當大的可能性是訓練程序將超出範圍並超出 w 的最佳值。在加權空間 w 中的這種現象稱為「振盪發散」，這將劣化系統性能的整個品質，且可能完全地破壞訓練程序。

上述的訓練程序將針對在訓練集合中的各個影像／詮釋對 $\{I^*, C^*\}$ 重覆執行。在學習模式中由系統處理的訓練資料愈多，則神經網路加權向量的各個分量值將收斂至最佳的數值，這些數值能滿足本發明中主控訓練處理的目的函數。應注意的是在訓練處理期間，無須執行射線搜尋演算法或者 Viterbi 演算法。

一旦訓練處理已產生一個適當的加權向量，系統可執行其辨識及給分工作，而無須進一步參考到訓練資料庫。此意即在某些實施例中，訓練可「在實驗室中」進行，而辨識和給分可「在現場」進行。現場使用的產品不需要提供給分資料庫或訓練演算法。在其他的實施例中，可能期望現場產品能執行再訓練或遞增訓練，在此狀況下，可能需要提供若干個儲存起來的選用訓練實例。

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明(41)

特別是，針對圖 1 2 中所示的「個人」用辨識器，藉由再訓練的過程，可使其適應單一使用者或少數使用者的差異性，此過程係根據使用者所提供的實例，如此可提升系統的性能。

當本發明中的方法和系統被實現為一個可攜式的手寫辨識裝置時，則位元對映的文字影像，數字串等已為使用者確認可用的資訊最好是存放在裝置中的非揮發性記憶體結構裡。這個記憶體結構的功能係存放著對應於影像／詮釋對 { I * / C * } 的位元對應及 A S C I I 格式資訊，在該辨識裝置使用了一段期間後，可由這些收集到的資訊中建立一個訓練資料集合。

當此訓練資料集合具有相當規模後，這個可攜式的裝置將可操作在「學習模式」。在此操作模式中，訓練資料集合中的影像會依據本發明的訓練過程被處理。在各個影像／詮釋對 { I * , C * } 被再處理後，則加權向量的各別分量將遞增性地予以調整，以求得上述的目的函數。

有許多本發明的其他實施例可容易地被建構起來。例如，不使用取自像素資訊的前置處理影像，系統的輸入可以是取自筆寫線條資訊的前置處理影像，或是取自筆劃資訊的列表（而非影像形式）。另一個例子是，系統的輸入可包括取自於聲頻信號，諸如，話語，的前置處理資訊。同理，可實現其他型式的輸出：輸出符號不僅可以是數字，亦可為阿拉伯字母，音位，全文字，抽象符號，或上述表示物的集合。很容易可以想像例如在一個具雜訊的通訊

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(42)

通道中傳輸之解碼及錯誤修正編碼符號的應用。

在本發明的另一個實施例中，由神經網路複合體執行的功能可利用任何具有以下能力的裝置來完成，這些能力是（1）接收一個輸入；（2）根據一組參數，產生一個可詮釋為一個得分或得分向量的輸出；（3）根據一個已知的微分向量，調整該組參數，此參數調整的方向將能使輸出朝著微分向量指定的方向改變。

在本發明的另一個實施例中，由“排列圖形”執行的功能可藉由一個傳統的動態可程式化格架，或者任何以需求方式處理串列資訊的裝置來完成，特別是：（1）接收描述可以是串列的一部份之各種實體的得分；（2）有效地辨識各種高得分的串列以及相對的詮釋；（3）有效地計算與一個已知的詮釋相符合之所有串列的總得分；以及（4）有效地計算其結果對輸入得分的靈敏度。

同時，在處理鏈中的模組數目可超過二個。各模組應具備（i）靈敏度輸出（若任何先前的模組具有可調的參數），（ii）靈敏度輸入（若這個模組或任何先前模組具有可調的參數），以及（iii）一般資料輸入以及資料輸出。

此處所討論的機率不需要在處理器和記憶體中以0和1之間的數字來表示。例如，最好將這些機率值以某一大的負數和0之間的對數機率值儲存起來，並且使用描述機率值串接及平行組合的計算步驟。

本發明中的方法和裝置可適用於詮釋表現在多種媒體

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明 (43)

中的輸入表象，例如電氣被動（圖形化的）式的記錄媒體，諸如紙張，毛皮，玻璃等；電氣主動式記錄媒體，諸如壓敏式書寫平面以及觸摸螢幕書寫及顯示平面；音訊記錄媒體，諸如人造或機具產生的語音；以及例如是空氣的媒體，此處揮動的筆劃經由編碼程序（藉由電氣主動式非接觸機制，例如 R F 位置感知，光學位置感知，電容式位置感知），再予以傳輸，記錄和／或使用本發明的系統和方法加以辨識。應注意的是，在後段提及的應用中，不需要將符號串圖形化地表示在平面上，而只要予以表象即可。

本發明中的系統和方法亦可適用於傳統的話語辨識系統。舉一個應用的實例，輸入的資料集合可以是以時域表示的一個記錄話語音調（亦即音位信號）。根據本發明，這個記錄下來的話語音調可分割為小的語音取樣（亦即語音細元），每一個取樣的間隔時間很短，各個話語細元將進行前置處理，且分割為話語細元，這些話語細元再結合起來構成「話語分段」，各個話語分段包含了代表至少在話語音調中一個音位的頻譜資訊。這些話語分段再結合起來形成串接分段，這些串接分段係利用本發明中的非循環圖形加以表示。之後，利用這些串接分段以及該組所有可能的音位串詮釋，本發明中的系統和方法則可針對最高得分音位串詮釋進行由果溯因機率值的計算。精通語音辨識技術的人士可輕易地明瞭語音辨識處理的細節。

可以理解的是，精通本技術領域的人士可容易地針對本發明的說明實施例作其他修改。而這些修改和變形必定

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

原

五、發明說明(44)

不超出本發明申請權利範圍所定義的範圍和精神。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

(原

四、中文發明摘要(發明之名稱：

使用新穎之由果溯因機率方式來對輸入表象作自動詮釋之系統與方法，以及最佳訓練之資訊處理網路

一種用以構成輸入表象之詮釋的方法和系統，其中的輸入表象係表現在一個媒體中，而詮釋是一串列的符號，其中的各個符號是一組已知符號中的一個符號，大體而言，此系統針對代表輸入表象的一組擷取輸入資料進行處理，如此可構成一組分段，這些分段再用來指定出一組串接分段，各個串接分段以及各個輸入表象的可能詮釋係表示在一個資料結構中，這個資料結構係利用一個具有排列成行列之二次元節點陣列的圖形所表示，這些行列係選擇性地由指向弧線連接起來，各條通過這些節點且沿著指向弧線伸展的路徑代表一個串接分段以及一個輸入表象的可能詮釋，所有的這些串接分段以及所有輸入表象的可能詮釋是由該組通過圖形的路徑來表示，對於圖形中的各列節點而言，將使用一種最佳化訓練神經資訊處理網路的複合體來針對已知的符號集合，產生一組得分。之後系統將針對一個或多個符號串詮釋，計算出一個由果溯因機率值，僅

(請先閱讀背面之注意事項再填寫本頁各欄)

英文發明摘要(發明之名稱：

SYSTEM AND METHOD FOR AUTOMATED INTERPRETATION
OF INPUT EXPRESSIONS USING NOVEL A POSTERIORI PROBABILITY
MEASURES AND OPTIMALLY TRAINED INFORMATION
PROCESSING NETWORKS

Abstract

A method and system for forming an interpretation of an input expression, where the input expression is expressed in a medium, the interpretation is a sequence of symbols, and each symbol is a symbol in a known symbol set. In general, the system processes an acquired input data set representative of the input expression, to form a set of segments, which are then used to specify a set of consegmentations. Each consegmentation and each possible interpretation for the input expression is represented in a data structure. The data structure is graphically representable by a graph comprising a two-dimensional array of nodes arranged in rows and columns and selectively connected by directed arcs. Each path, extending through the nodes and along the directed arcs, represents one

四、中文發明摘要 (發明之名稱：)

透過對擷取輸入資料集合作分析，可求出各個由果溯因機率值，進而針對競爭的輸入表象詮釋，取得高可靠度的機率值，本發明可適用於任意長度的手寫字符串，且可輕易地應用在語音識別系統中。

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

英文發明摘要 (發明之名稱：)

consegmentation and one possible interpretation for the input expression. All of the consegmentations and all of the possible interpretations for the input expression are represented by the set of paths extending through the graph. For each row of nodes in the graph, a set of scores is produced for the known symbol set, using a complex of optimally trained neural information processing networks. Thereafter the system computes an a posteriori probability for one or more symbol sequence interpretations. By deriving each a posteriori probability solely through analysis of the acquired input data set, highly reliable probabilities are produced for competing interpretations for the input expression. The principles of the present invention can be practiced with cursively written character-strings of arbitrary length and can be readily adapted for use in speech recognition systems.

訂

線

六、申請專利範圍

附件二A:

第84101589號專利申請案

中文申請專利範圍修正本

民國87年8月修正

1. 一種用以分析一個輸入表象並且針對此輸入表象之可能詮釋進行給分的系統，其包含：

分段產生裝置，用以分析代表該個輸入表象的一組輸入資料並且將這組輸入資料區分為若干個分段，其中的各個分段具有可指定的邊界且可予以歸類成可能代表在一預設符號集合中若干符號中的任一個符號；

該輸入資料集合包括一組像素，其與一由一串列之以圖像表示的符號之被擷取影像相關，及其中該分段產生裝置對該組像素進行分析，並且將該組像素區分為若干個影像分段，如此使得上述的各影像分段具有指定的邊界且可歸類成可能代表存在於一預設的字符集合中之上述若干字符中任何一個或多個的字符；

分段給分裝置，用以分析上述若干分段中的各個分段，並且將一個得分指定給與上述預設符號集合中的特定符號相關之分段中的各個可能分類；

表示裝置，用以代表上述輸入表象的數個可能詮釋，以及代表數個影像串接分段，其中各個可能詮釋包含上述符號的不同組合串列，而各個上述的串接分段包含上述分段的不同組合串列；

串接分段給分裝置，用以將得分依據指定給上述分段

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

煩請委員明示87年8月7日所提之修正本有無變更實質內容是否准予修正。

六、申請專利範圍

附件二A:

第84101589號專利申請案

中文申請專利範圍修正本

民國87年8月修正

1. 一種用以分析一個輸入表象並且針對此輸入表象之可能詮釋進行給分的系統，其包含：

分段產生裝置，用以分析代表該個輸入表象的一組輸入資料並且將這組輸入資料區分為若干個分段，其中的各個分段具有可指定的邊界且可予以歸類成可能代表在一預設符號集合中若干符號中的任一個符號；

該輸入資料集合包括一組像素，其與一由一串列之以圖像表示的符號之被擷取影像相關，及其中該分段產生裝置對該組像素進行分析，並且將該組像素區分為若干個影像分段，如此使得上述的各影像分段具有指定的邊界且可歸類成可能代表存在於一預設的字符集合中之上述若干字符中任何一個或多個的字符；

分段給分裝置，用以分析上述若干分段中的各個分段，並且將一個得分指定給與上述預設符號集合中的特定符號相關之分段中的各個可能分類；

表示裝置，用以代表上述輸入表象的數個可能詮釋，以及代表數個影像串接分段，其中各個可能詮釋包含上述符號的不同組合串列，而各個上述的串接分段包含上述分段的不同組合串列；

串接分段給分裝置，用以將得分依據指定給上述分段

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

煩請委員明示87年8月7日所提之修正本有無變更實質內容是否准予修正。

六、申請專利範圍

的得分指定給上述若干個串接分段；

候選詮釋辨識裝置，用以依據指定給上述若干分段之得分，自上述若干個可能詮釋中，辨識出一個或多個候選符號詮釋；

符號串列給分裝置，用以依據指定給上述若干分段中一個或多個分段的得分，將分數指定給上述一個或多個候選詮釋；

第 1 得分評估裝置，用以評估出指定給一個或多個候選詮釋的得分；

第 2 得分評估裝置，用以評估指定給上述若干可能詮釋的得分；

經正規化的得分產生裝置，其藉由使用的是上述若干可能詮釋的評估得分，來針對各候選詮釋，產生一個經正規化的得分。

2. 如申請專利範圍第 1 項中的系統，其中該分段給分裝置針對上述若干影像分段中的各個影像分段進行分析，並將一個得分指定給上述影像分段中各個可能的分類，其中各個經指定的得分係與上述預設字符集合中某一特定字符相關。

3. 如申請專利範圍第 2 項中的系統，其中該表示裝置代表若干個可能的字符串以及若干個影像串接分段，其中各個可能字符串包含一串列上述的字符，而各個上述的串接分段包含一串列上述的影像分段。

4. 如申請專利範圍第 3 項中的系統，其中的串接分

六、申請專利範圍

段給分裝置依據指定給影像分段的得分，將得分指定給上述若干個影像串接分段，且其中該候選字符串列辨識裝置依據指定給該影像分段的得分，辨識出一個或多個候選字符串列。

5. 如申請專利範圍第4項中的系統，其中該符號串列給分裝置依據指定給影像分段的得分，將得分指定給上述一個或多個的候選字符串列，及其中該第1給分評估裝置對指定給該一個或更多個候選字符串的得分進行評估。

6. 如申請專利範圍第5項中的系統，其中的第2得分評估裝置針對指定給上述多個可能字符串的得分進行評估，及該得分正規化裝置使用上述多個可能字符串的評估得分來對指定給上述各候選字符串的得分進行正規化。

7. 如申請專利範圍第6項中的系統，其中上述的表示裝置包含一個可用某一特定圖形來表示的資料結構，而這個圖形包括一個以列行排列的二次元節點陣列，此陣列係選擇性地以指向弧線連接起來，其中各行的節點係利用一個字符位置予以標記，而各列節點則以一個上述的影像分段予以標記，標記的次序則對應至該擷取影像的空間結構，且其中通過這些節點且沿著該指向弧線伸展的各條路徑代表一個上述的串接分段以及一個上述的可能字符串，而且實質上所有上述的影像串接分段和上述所有的可能字符串是由該組通過上述圖形的路徑來表示。

8. 如申請專利範圍第7項中的系統，其中的各個節點亦包含一組辨識弧線，而各個辨識弧線代表一個上述的

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

六、申請專利範圍

字符，並且與一個被指定的得分相關。

9. 如申請專利範圍第1項中的系統，其中的表示裝置隱含性地代表上述若干個可能詮釋以及上述若干個影像串接分段。

10. 如申請專利範圍第9項中的系統，其中的表示裝置包含一種可以一圖形表示的資料結構，此圖形包含一個排列成行列的二次元節點陣列，此陣列選擇性地被指向弧線連接起來，其中的各行節點可以一個符號位置予以標記，且其中的各列節點可以上述的一個分段予以標記，而標記的次序實質上對應上述輸入資料集合的序列結構，及其中各條通過節點且沿指向弧線伸展的路徑代表一個上述的串接分段以及一個輸入表象的可能詮釋，實質上所有的串接分段以及所有的可能詮釋可以該組通過該圖形的路徑來表示。

11. 如申請專利範圍第1項中的系統，其中的表示裝置包含一個可以某一圖形代表的資料結構，這個圖形包含一個排列成行列的二次元節點陣列，此陣列選擇性地被指向弧線連接起來，其中各行節點係利用一個符號位置予以標記，而各列節點係利用一個分段予以標記，標記的次序實質上對應於該組輸入資料的串列結構，及其中通過這些節點且沿著指向弧線伸展的各條路徑代表了一個串接分段以及該輸入表象的一可能詮釋，且所有的串接分段以及所有的可能詮釋是由該組通過上述圖形的路徑所表示。

12. 一種用以構成一輸入表象之詮釋的方法，其中

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

六、申請專利範圍

的輸入表象可顯現在一個媒體中，而詮釋是一串列的符號，且各個符號是一個預設符號集合中的一個元素，這種方法包含以下的步驟：

(a) 擷取一組代表該輸入表象的輸入資料，該組輸入資料包括一組像素，其與一由串列之以影像表示的符號之被擷取的影像相關連；

(b) 對此組輸入資料進行處理以構成一組分段，其中的各個分段至少是該組經擷取的輸入資料的部份子集合，且可予以歸類成代表上述預設符號集合中的任何一個符號；

(c) 構成一個代表一組串接分段以及該輸入表象之一組可能詮釋的資料結構，其中的各串接分段包含一組集合起來以表示該組輸入資料的分段，並以一次序加以排列且該次序實質上保留了該組輸入資料的串列結構，上述輸入表象的各個可能詮釋包含一個可能符號串列，而該可能符號串中的各個符號係選自於上述上述的預設符號集合且佔有一個上述可能符號串中的符號位置，該資料結構係利用一個圖形加以表示，這個圖形包含一個排列成行列的二次元節點陣列，此陣列選擇性地被指向弧線連接起來，其中的各行節點可利用一個符號位置予以標記，各列節點可以一個影像分段予以標記，標記的次序係對應上述擷取輸入資料集合的邏輯結構，而各條通過這些節點且沿著這些指向弧線伸展的路徑代表一個串接分段以及一個輸入表象的可能詮釋，所有的串接分段和所有輸入表象的可能詮釋

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

六、申請專利範圍

係由該組通過圖形的路徑來表示；

(d) 針對該圖形中的各列節點，產生一組預設符號集合所使用的得分，其中的預設符號集合是以該列的各節點來表示，而該組得分中的每一個得分產生包括了分析用以標記一列節點的分段，而上述之得分即針對這些節點所產生；

(e) 隱含性或外顯地將一個路徑得分歸屬到通過圖形的路徑；以及

(f) 分析在步驟(e)中歸屬至通過圖形之路徑的路徑得分，以便針對該輸入表象，選擇一個或多個可能詮釋。

13. 如申請專利範圍第12項中的方法，其中的各個節點亦包含一組辨識弧線，而其中的各個辨識弧線代表一個上述預設的符號並且與一個在步驟(d)中產生的得分相關。

14. 如申請專利範圍第13項中的方法，其中的步驟(d)包含了使用若干可調參數以產生該組得分的步驟。

15. 如申請專利範圍第14項中的方法，其中一種資訊處理裝置的特性可由上述若干個可調整參數予以確定，而此種裝置係使用在步驟(d)中，以便分析各個分段且產生其中的該組得分。

16. 如申請專利範圍第13項中的方法，其中的步驟(f)包含：針對上述輸入表象的至少一個可能詮釋，

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

六、申請專利範圍

計算出一個對應於一個由果溯因機率的數量，其中計算得出的每一個數量是一個具有分子部份和分母部份的比值，其中的分子部份對應至路徑得分的總和，而所謂的路徑得分係針對實質之所有通過代表該輸入表象的該可能詮釋之圖形的路徑而言，上述的各個路徑得分對應至與上述路徑的辨識弧線相關之得分的乘積；分母部份對應路徑得分的總和，而這些路徑得分是對實質上所有通過圖形的路徑而言，這個圖形代表實質上所有輸入表象的所有可能詮釋，其中的各個路徑得分對應沿一條路徑伸展之辨識弧線相關之得分的乘積。

1 7 . 如申請專利範圍第 1 6 項中的方法，在步驟 (f) 中亦包括：

(1) 確定通過具有最高路徑得分之圖形的路徑，

(2) 辨識出輸入表象的可能詮釋，而其中的輸入表象係由子步驟 (1) 中確定的路徑所表示，

(3) 針對子步驟 (2) 中辨識出之輸入表象的可能詮釋，計算其數量；以及

(4) 將子步驟 (3) 中計算求得的數量以及代表輸入表象之可能詮釋的標記提供作為輸出，而這個可能詮釋係於子步驟 (2) 中辨識而得。

1 8 . 如申請專利範圍第 1 6 項中的方法，在步驟 (f) 中亦包含下列步驟：

(1) 確定一組通過具有一組高路徑得分值之圖形的路徑，

六、申請專利範圍

(2) 辨識一組輸入表象的可能詮釋，其由子步驟(1)中確定的該組路徑所表示，

(3) 針對在子步驟(2)中辨識出來的輸入表象可能詮釋集合，計算一組數量；

(4) 分析該組在子步驟(3)中計算得出的數量，以確定該輸入表象的那一個可能詮釋具有一個高得分的由果溯因機率；以及

(5) 將子步驟(2)中辨識出的輸入表象之可能詮釋以及代表高得分由果溯因機率值的標記，提供作為輸出，其中的機率值是在子步驟(4)中所求得。

19. 如申請專利範圍第16項中的方法，其中計算求得的各個由果溯因機率值是一個具有分子部份和分母部份的比值，其中的步驟(f)亦包含：

(1) 確定一組通過具有一組高路徑得分之圖形的路徑，

(2) 辨識出輸入表象的一組可能詮釋，它們是由子步驟(1)中確定的一組路徑來表示，

(3) 針對輸入表象在子步驟(2)中辨識而得之該組可能詮釋，計算出一組數量；以及

(4) 將子步驟(2)中辨識出的該輸入表象之可能詮釋以及在子步驟(3)中計算所得的數量，提供作為輸出。

20. 如申請專利範圍第14項中的方法，在步驟(d)中，該組可調參數指定出該分段與該組得分之間的關

六、申請專利範圍

係，其中的分段是提供給資訊處理裝置，而該組得分則是由該資訊處理裝置所產生。

2 1 . 如申請專利範圍第 2 0 項中的方法，亦包含訓練該資訊處理裝置的程序，步驟為：

(1) 使用上述資訊處理裝置來處理數個已知符號串，以及

(2) 對各個已知的串列，遞增性地調整該組可調參數，使得指定給正確詮釋的機率值在平均數值上能有所增加，而指定給不正確詮釋的機率值在平均值上能有所減少。

2 2 . 如申請專利範圍第 2 1 項中的方法，其中的資訊處理裝置包含一個神經資訊處理網路。

2 3 . 如申請專利範圍第 1 2 項中的方法，其中的輸入表象是以印刷或手寫方式來表達，且以圖形記錄在一個記錄媒體中。

2 4 . 一種用構成輸入表象之一個詮釋的系統，其中的輸入表象係表達在一個媒體中，該詮釋是一串列的符號，且其中的各個符號是在一組預設符號集合中的一個元素，此系統包含：

資料集合擷取裝置，用以擷取表示該輸入表象之輸入資料集合；

資訊處理裝置，用以處理上述擷取到的資料集合，以產生若干個分段，各個分段具有可指定的邊界，且儘可能歸類成代表在一預設符號集合中若干符號的任一個符號；

六、申請專利範圍

串接分段指定裝置，用以產生指定一組串接分段之資料，其中各個串接分段包含一組分段，這些分段合起來代表取得的輸入資料集合且排列的次序實質上保留著上述取得輸入資料集合的串列結構；

符號串詮釋指定裝置，用以提供指定一組輸入表象之可能詮釋的資料，各個輸入表象的可能詮釋包含一個可能符號串，可這個可能符號串中的各個符號是選自該預設符號集合，且在該可能符號串中，佔據一個符號位置；

資料儲存裝置，用以存放資料結構，所產生的資料代表各個串接分段以及輸入表象的可能詮釋，其中的資料結構係以一個具二次元節點陣列的圖形來表示，這個陣列排列成行與列，且利用指向弧線選擇性地予以連接，而其中的各行節點可利用一個符號位置予以標記，各列節點則可以一個分段來予以標記，標記的次序係對應擷取輸入資料集合的串列結構，其中各條通過這些節點且沿著指向弧線伸展的路徑代表一個串接分段以及一個輸入表象的可能詮釋，其中的該組串接分段以及該組輸入表象的可能詮釋可利用一組通過該圖形的路徑來代表；

料分段分析裝，用以分析各分段之資置，且針對圖形中的各列節點，產生一組該符號集合所使用的得分，這組符號可由該列節點中的各個節點來表示；

路徑得分計算裝置，用以針對通過該圖形的各條路徑，計算出一個路徑得分；以及

路徑得分分析裝置，用以分析計算得出的路徑得分，

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

六、申請專利範圍

以便選取一個或多個輸入表象的可能詮釋。

25. 如申請專利範圍第24項中的系統，其中各個節點包含一組辨識弧線，而各個辨識弧線代表一個已知的符號且與其中一個計算求得的得分相關。

26. 如申請專利範圍第25項中的系統，其中的路徑得分分析裝置亦包含一種對應一由果溯因機率值而計算出一數量的裝置，而上述的由果溯因機率值係針對每一個輸入表象的可能詮釋而言。

27. 如申請專利範圍第26項中的系統，其中計算求出的各個數量係一個具有分子部份及分母部份的比值，

其中的分子部份對應實質上所有通過代表該輸入表象的一可能詮釋之圖形之路徑的路徑得分總和，而其中的各個路徑得分對應於沿一該路徑伸展的辨識弧線相關之得分的乘積，及

其中的分母部分對應於實質上所有通過代表該輸入表象之實質上所有可能的詮釋之圖形的路徑之路徑得分的總和，其中的各個路徑得分對應於沿一該路徑伸展之辨識於相關之得分的乘積。

28. 如申請專利範圍第24項中的系統，其亦包含

:

用以確定通過具有最高路徑得分之圖形的路徑之裝置

,

用以辨識出輸入表象之可能詮釋的裝置，這個可能詮釋係以確定所得之具最高路徑得分的路徑來代表，

六、申請專利範圍

用以針對輸入表象之各個可能詮釋，以計算出該數量的裝置；以及

用以將代表計算所得出之數量和輸入表象之可能詮釋，提供作為輸出的裝置。

29. 如申請專利範圍第28項中的系統，其中的路徑得分分析裝置亦包含

用以確定一組路徑的裝置，這一組路徑須通過具有一組最高路徑得分的圖形，

用以辨識出一組輸入表象之可能詮釋的裝置，其中的輸入表象是由確定出的該組路徑所表示，

針對輸入表象所辨識而得的該組可能詮釋，以計算出一組數量的裝置，

一種裝置，用以分析上述計算得到的該組數量，以及確定那一個輸入表象的可能詮釋具有該組最高路徑得分的最高之由果溯因機率值；以及

一種將代表輸入表象之可能詮釋的標記和所確定出的最高之由果溯因機率值提供作為輸出的裝置，上述的輸入表象具有最高的由果溯因機率值。

30. 如申請專利範圍第26項中的系統，其中的分段分析裝置具有一組可調參數，其能指定提供作分析用之分段和由該分段分析裝置輸出之一組得分之間的關係。

31. 如申請專利範圍第30項中的系統，亦包含用以訓練該系統的系統訓練裝置，此裝置使用若干個訓練資料集合，各個訓練資料集合包含一組輸入表象的取得資料

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

六、申請專利範圍

集合以及該個輸入表象的已知正確詮釋，該系統訓練裝置亦包括參數調整裝置，此裝置可遞增地調整該組可調參數，如此能夠增加該已知正確詮釋的平均機率量度，同時降低該組已知不正確詮釋的平均機率量度。

3 2. 如申請專利範圍第 2 6 項中的系統，其中的輸入表象係以圖形型式記錄在一個記錄媒體中，

3 3. 如申請專利範圍第 3 2 項中的系統，其中的輸入表象係以印刷或手寫方式表達。

3 4. 一種用以形成一輸入表象之詮釋的系統，其中的輸入表象是表達在一個媒體中，該詮釋是一串的符號，而其中的各個符號就是一個預設符號集合中的元素，此系統包括：

影像擷取裝置，用以擷取該輸入表象之一個影像；

影像處理裝置，用以處理該影像以構成一組影像分段，其中的各個影像分段是該擷取影像的一個子影像；

影像串接分段指定裝置，用以產生指定一組影像串接分段之資料，其中各個串接影像分段包含一組影像分段，這些影像分段集合起來代表該個擷取影像，且排列的次序實質上保留了該擷取影像的空間結構；

符號串詮釋指定裝置，用以產生可指定一組輸入表象之可能詮釋的資料，其中的各個輸入表象的可能詮釋包括一串列的符號，在此串列符號中的各個符號係選自於該預設的符號集合，且在該串列符號中佔有一個符號位置；

資料儲存裝置，用以存放資料結構，所產生的資料是

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

六、申請專利範圍

代表著各個影像串接分段和各個輸入表象的可能詮釋，其中各個資料結構是以一個指向非循環圖形來表示，而這個圖形包括一個排列成行列的二次元節點陣列，且選擇性地被指定弧線連接起來，其中的各行節點可以一個符號位置予以標記，各列節點可以一個影像分段予以標記，其排列的次序係對應該擷取影像的空間結構，而其中各條通過這些節點且沿這些指向弧線伸展的路徑代表其中一個影像串接分段以及一個輸入表象的可能詮釋，而所有的影像串接分段以及所有輸入表象的可能詮釋均由通過圖形中的該組路徑來表示；

影像分段分析裝置，用以分析各影像分段，針對圖形中的各列節點，該裝置產生一組用於該預設符號集合的得分，而這個預設的符號集合是由該列中的各個節點來表示；

路徑得分計算裝置，用以針對通過圖形中的各條路徑，計算一個路徑得分；以及

路徑得分分析裝置，用以分析上述計算得出之路徑得分，如此可針對輸入表象，選出一個或多個可能詮釋。

35. 如申請專利範圍第34項中的系統，其中的輸入表象係以圖形型式記錄在一個電氣被動式媒體中。

36. 如申請專利範圍第34項中的系統，其中的輸入表象係以圖形型式記錄在一個電氣主動式媒體中。

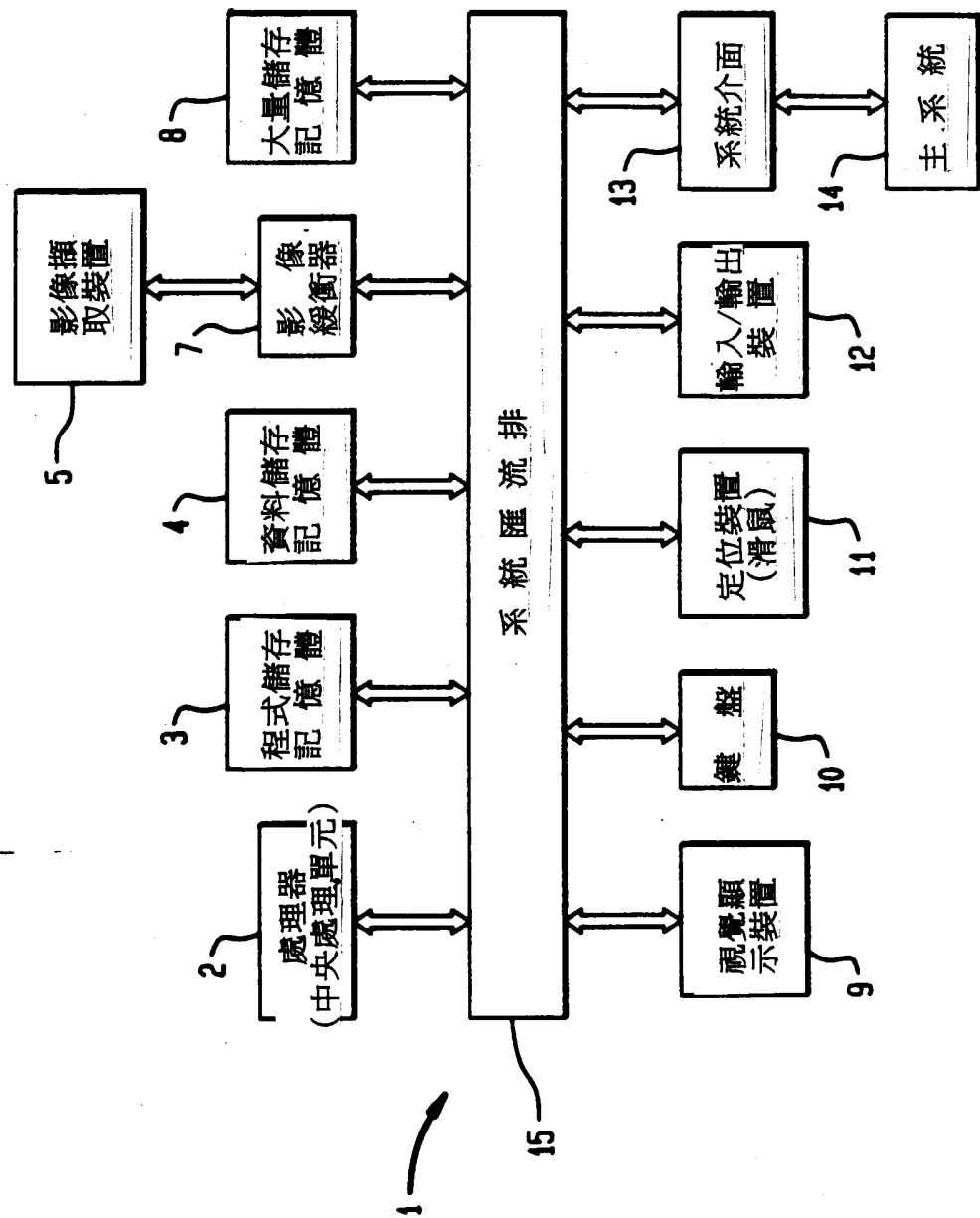
37. 如申請專利範圍第34項中的系統，其中的輸入表象是以印刷或手寫方式書寫而成。

8401589

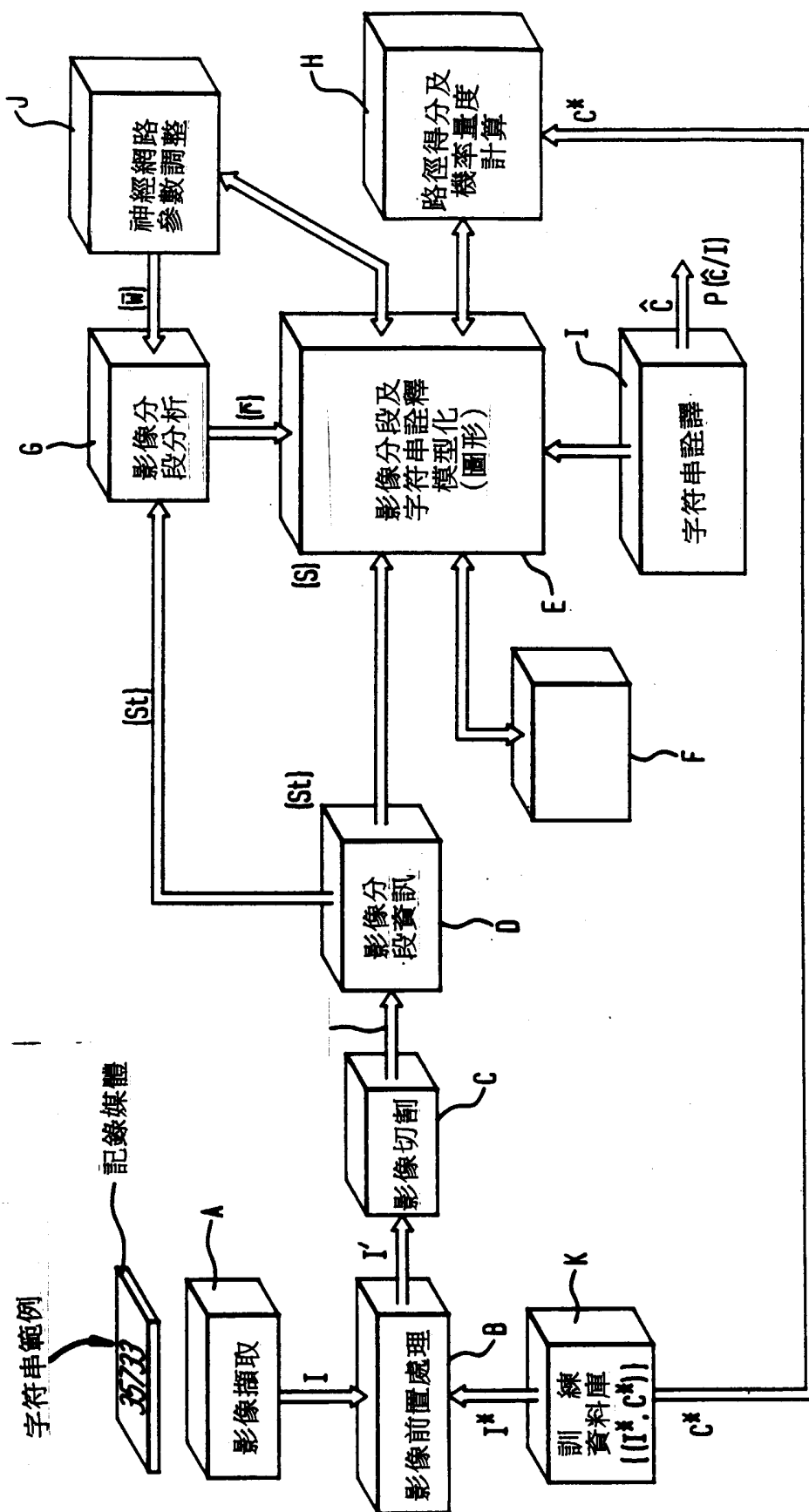
1/15

821700

第 1 圖



第 2 圖



第 3 圖

35733

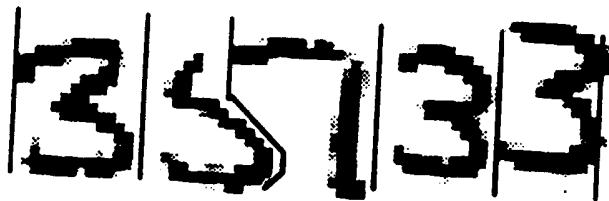
第 4A 圖

35733

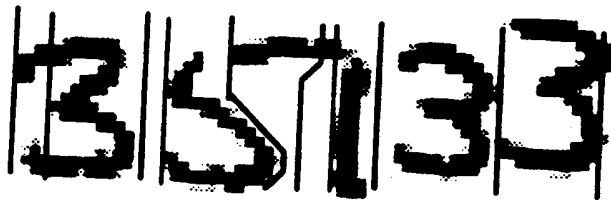
第 4B 圖

3 5 7 3 3

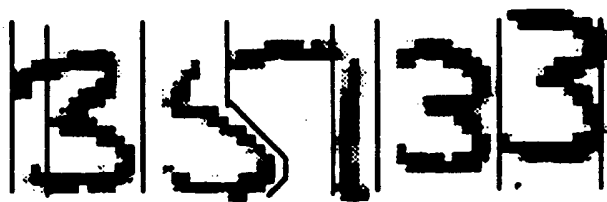
第 4C 圖



第 4D 圖



第 4E 圖



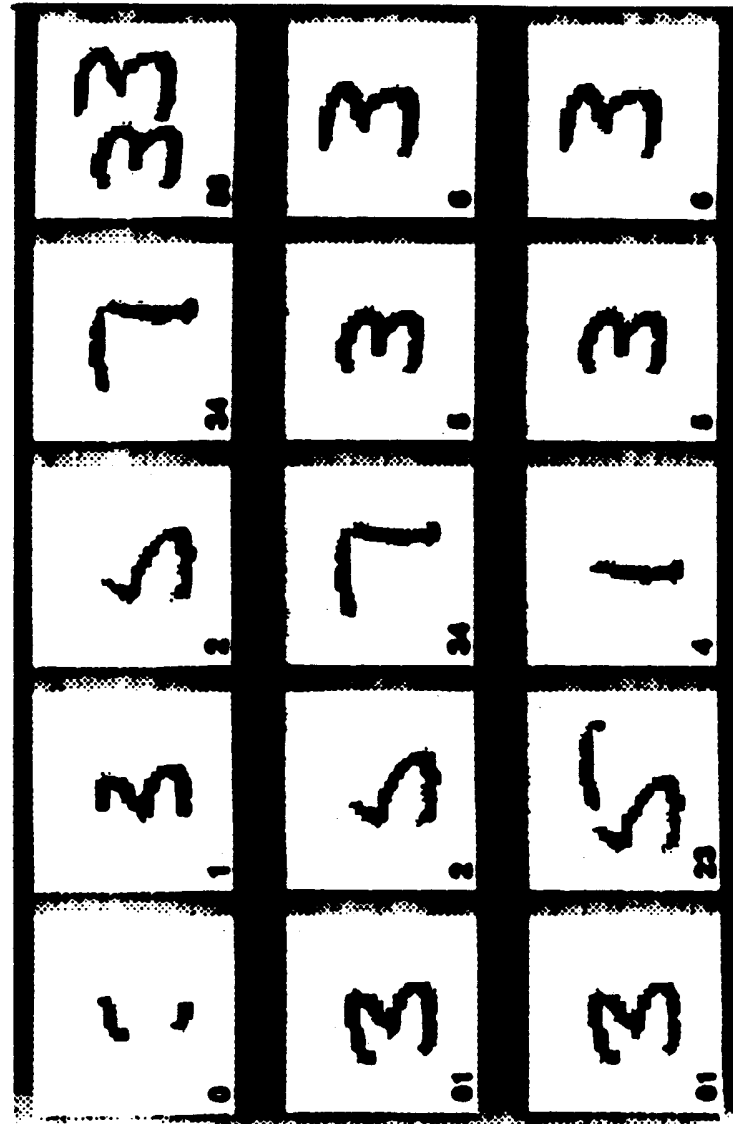
第 5 圖

0 7	1 3	2 5	3 —
4 1	5 3	6 3	

第 6 圖

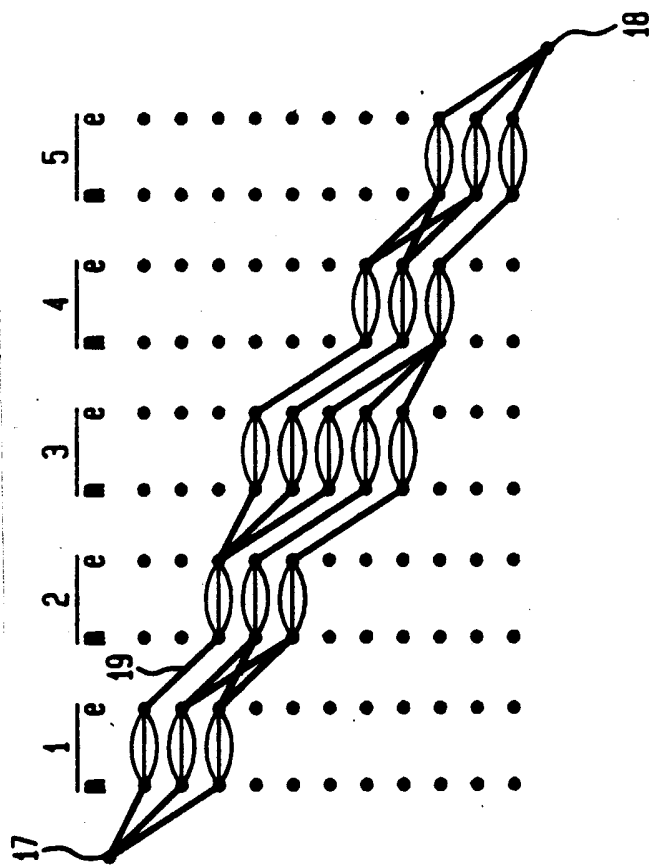
0 7	01 3	1 3	2 5
23 5	24 57	25 7	4 1
3 3	36 33	6 3	

第7圖



第 8 圖

字 符 位 置



影像左緣

0 01 1 2 23 234 34 4 5 56 6

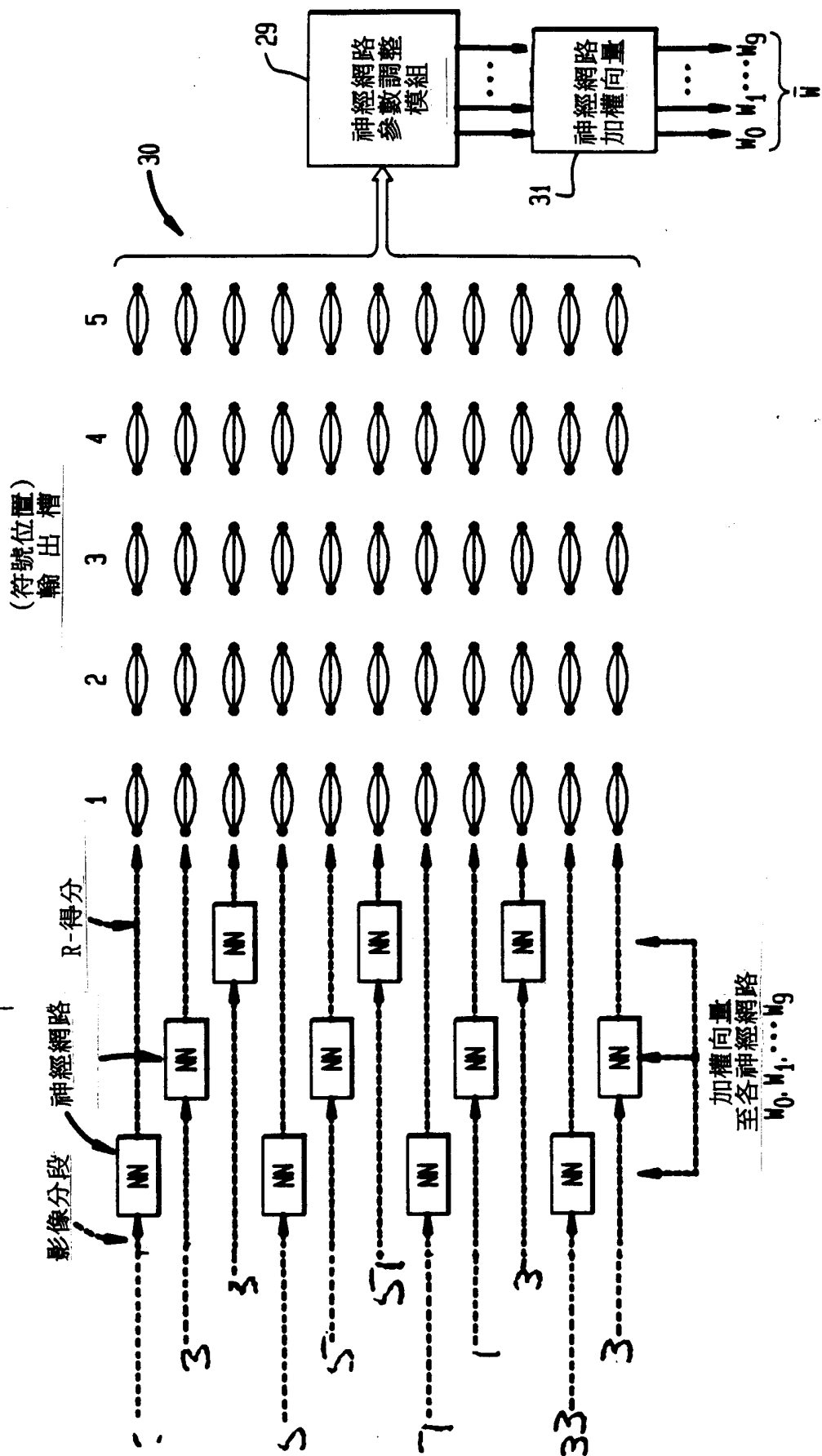
影像右緣

分段

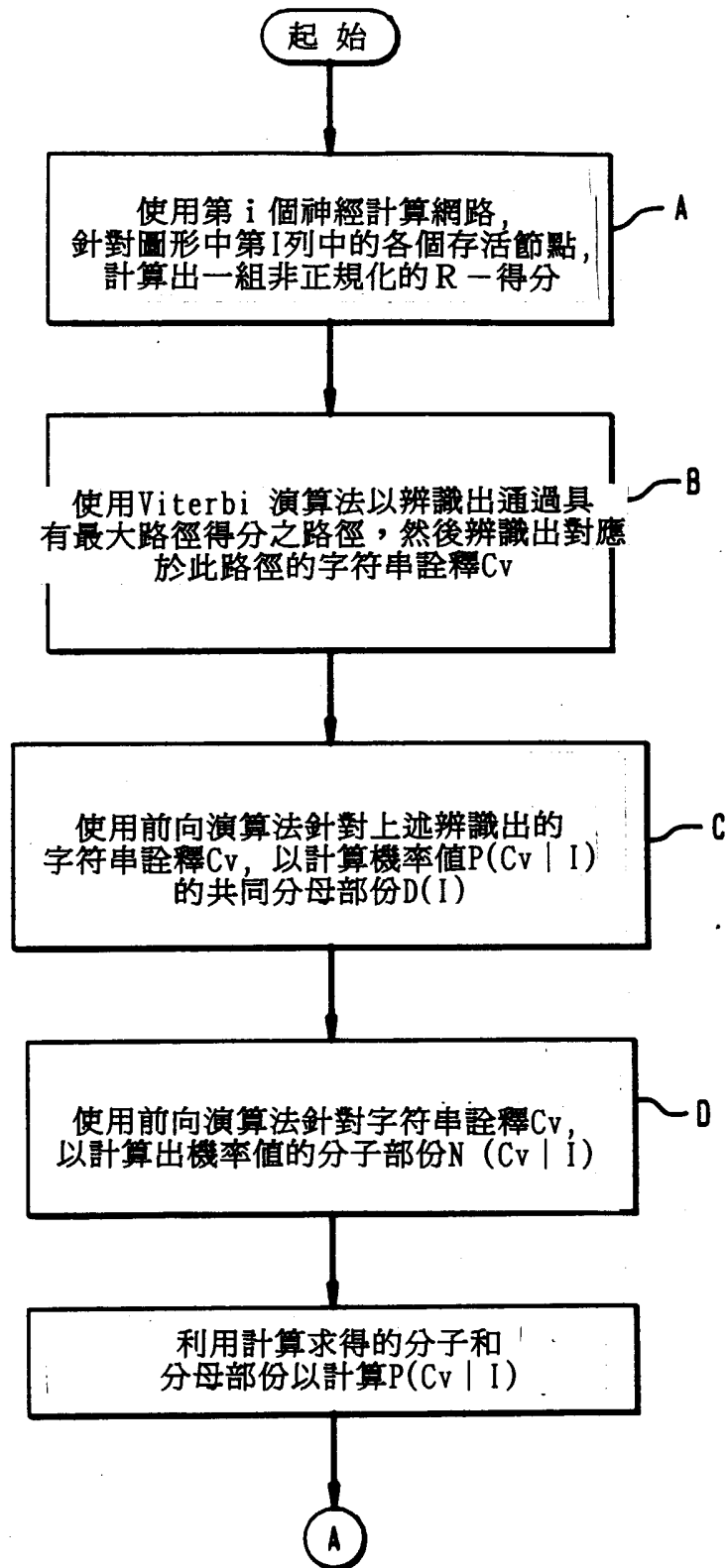
影像分段Si

3 3 5 7 33 3 5 1 3

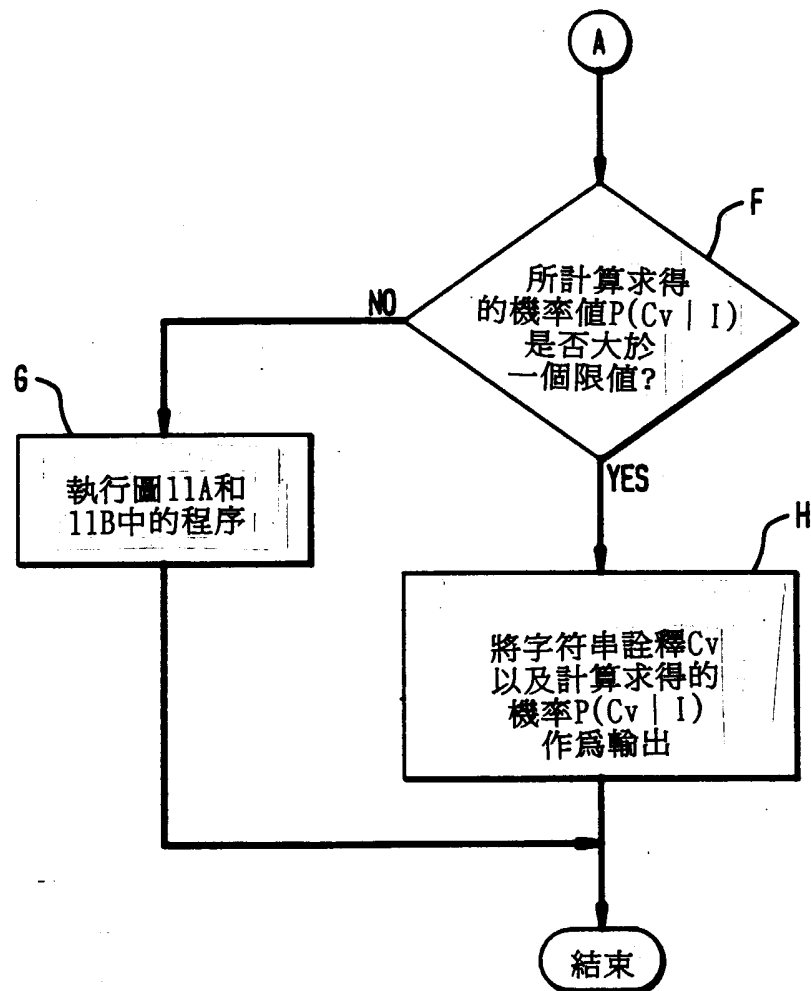
第 9 圖



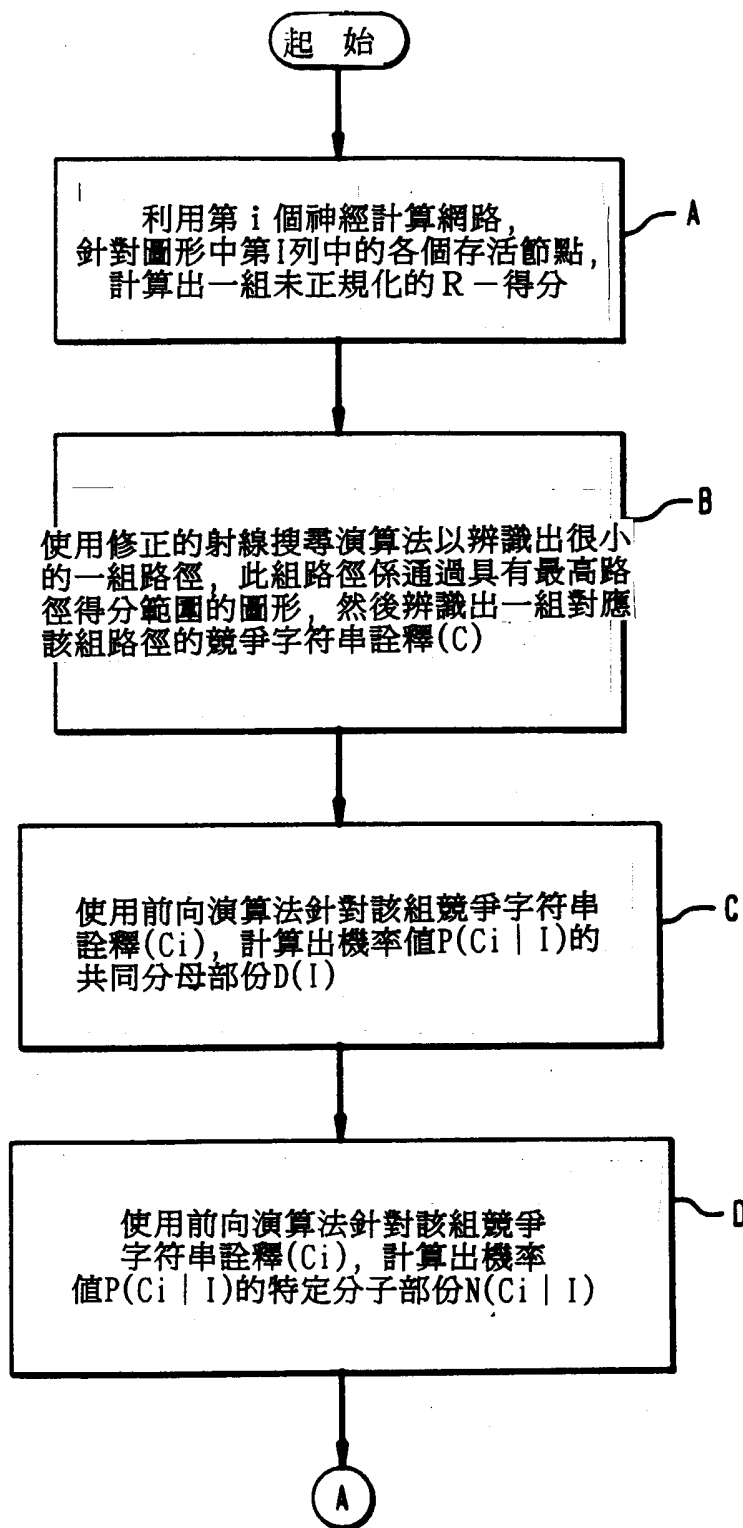
第 10A 圖



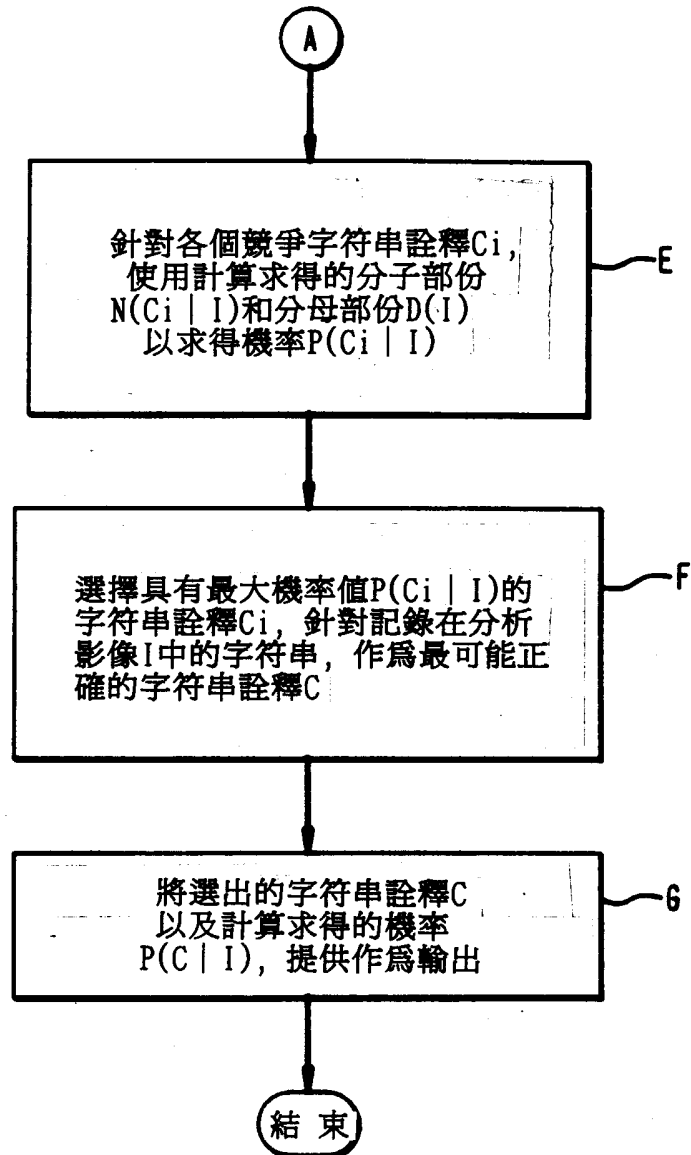
第 10B 圖



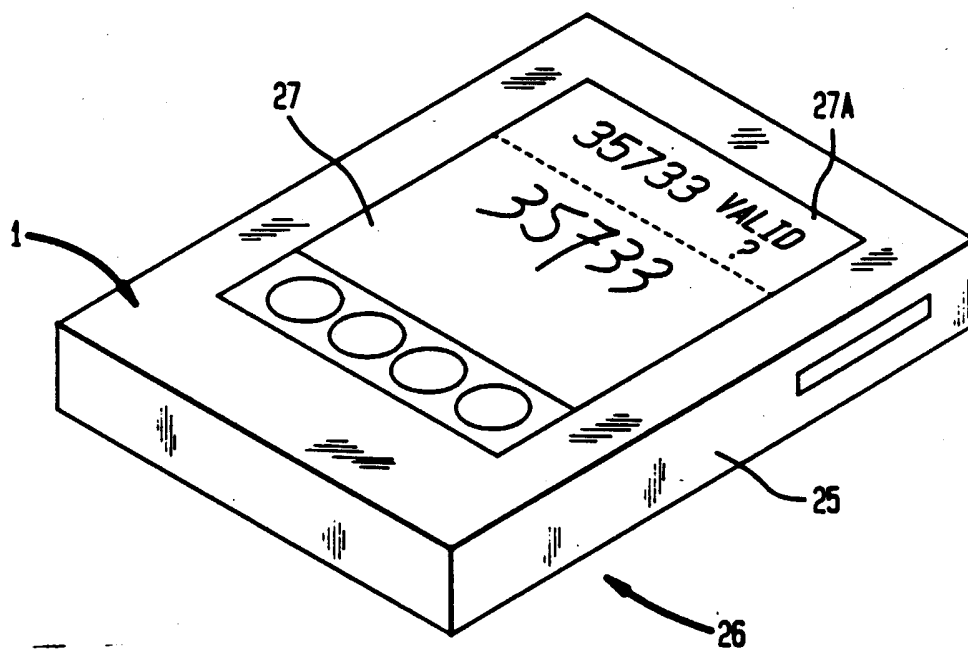
第 11A 圖



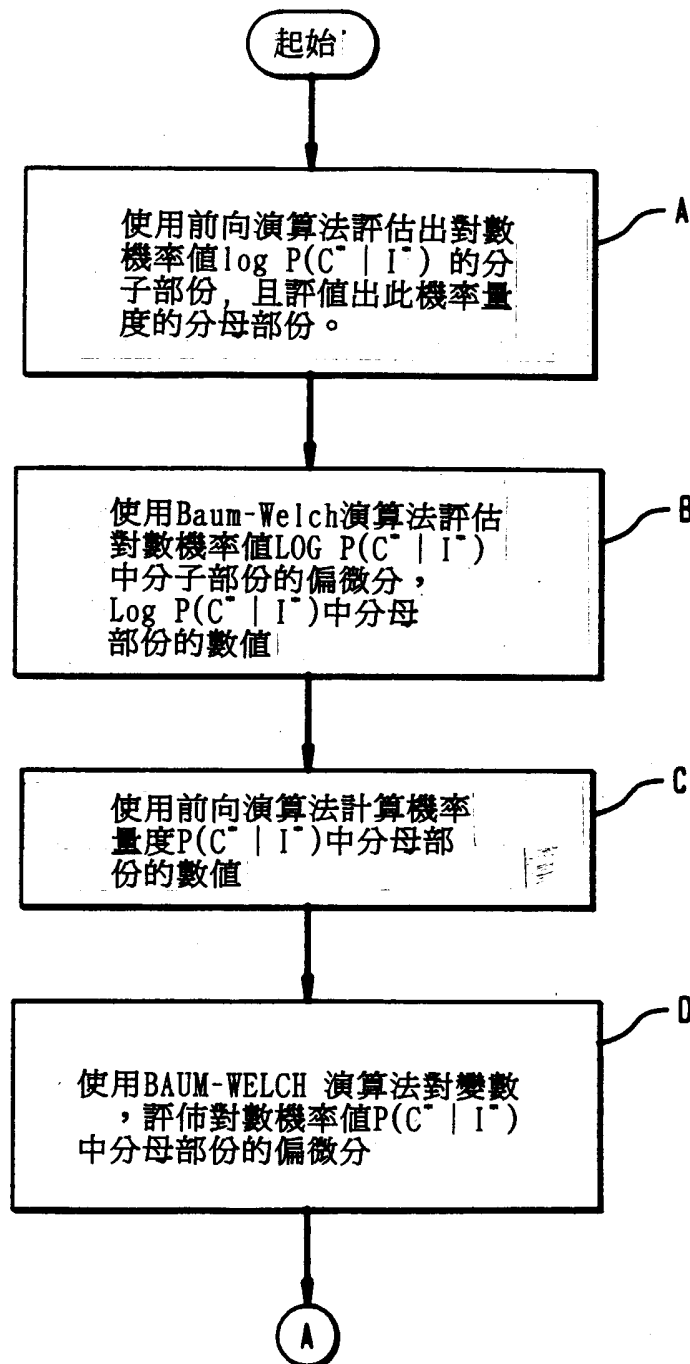
第 11B 圖



第 12 圖



第13A圖



第13B圖

