



(21)申請案號：104126837

(22)申請日：中華民國 104 (2015) 年 08 月 18 日

(51)Int. Cl. : G06N7/02 (2006.01)

(71)申請人：國立交通大學(中華民國) NATIONAL CHIAO TUNG UNIVERSITY (TW)
新竹市東區大學路 1001 號

(72)發明人：林進燈 LIN, CHIN TENG (TW)；邱晴瑜 CHING-YU CHIU (TW)；吳尚林 WU, SHANG LIN (TW)

(74)代理人：林火泉

申請實體審查：有 申請專利範圍項數：12 項 圖式數：3 共 12 頁

(54)名稱

遷移學習演算方法

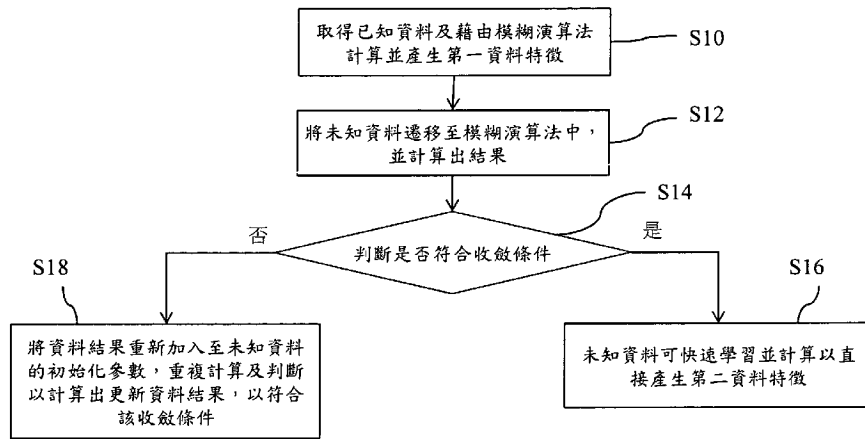
CALCULATION METHOD OF LEARNING MIGRATION

(57)摘要

本發明係揭露一種遷移學習演算方法，首先取得一筆已知資料及隨機產生第一初始化參數，並藉由一模糊演算法計算已知資料以獲得第一資料特徵，再將一筆未知資料遷移至模糊演算法中，並以第一筆資料特徵作為未知資料的初始化參數計算出資料結果，若資料結果收斂則表示未知資料可以快速學習已知資料的資料結果計算，以直接產生一第二資料特徵，若不收斂則須將資料結果重新加入初始化參數中，重複計算以滿足收斂條件。本發明可以使一筆新的資料快速經既有資料的特徵以計算，除了省去資料的計算時間，並達成計算完資料結果之準確性。

The present invention discloses a calculation method of learning migration, wherein obtain a known data and randomly generates a first initialization parameters, and calculate the known data to obtain the first information feature by a fuzzy algorithm, then moved an unknown data into the fuzzy algorithm and calculated data results by the first information feature as the unknown data's first initialization parameters. If the data result is convergence, the unknown data can quickly learn the known data's result to calculate in order to generate a second information feature. If the data result is not convergence, the data results will be re-added to the initialization parameters that repeat and repeat calculations to meet the convergence condition. The present invention can quickly calculate a new information by existing data feature, not only to save the computing time but also to achieve accuracy of the data result.

指定代表圖：



第二圖

201709115

【發明摘要】

【中文發明名稱】遷移學習演算方法

【英文發明名稱】calculation method of learning migration

【中文】

本發明係揭露一種遷移學習演算方法，首先取得一筆已知資料及隨機產生第一初始化參數，並藉由一模糊演算法計算已知資料以獲得第一資料特徵，再將一筆未知資料遷移至模糊演算法中，並以第一筆資料特徵作為未知資料的初始化參數計算出資料結果，若資料結果收斂則表示未知資料可以快速學習已知資料的資料結果計算，以直接產生一第二資料特徵，若不收斂則須將資料結果重新加入初始化參數中，重複計算以滿足收斂條件。本發明可以使一筆新的資料快速經既有資料的特徵以計算，除了省去資料的計算時間，並達成計算完資料結果之準確性。

【英文】

The present invention discloses a calculation method of learning migration, wherein obtain a known data and randomly generates a first initialization parameters, and calculate the known data to obtain the first information feature by a fuzzy algorithm, then moved an unknown data into the fuzzy algorithm and calculated data results by the first information feature as the unknown data's first initialization parameters. If the data result is convergence, the unknown data can quickly learn the known data's result to calculate in order to generate a second information feature. If the data result is not convergence, the data results will be re-added to the initialization parameters that repeat and repeat calculations to meet the convergence condition. The present invention can quickly calculate a new information by existing data feature, not only to save the computing time but also to achieve accuracy of the data result.

FIG. 1

【指定代表圖】：第（二）圖。

【代表圖之符號簡單說明】

無

【發明說明書】

【中文發明名稱】 遷移學習演算方法

【英文發明名稱】 calculation method of learning migration

【技術領域】

【0001】 本發明係一種用於不確定性資料的分類方法，特別是一種藉由模糊演算法作學習的遷移學習演算方法。

【先前技術】

【0002】 隨著網際網路的發達以及科技的進步，無論是上網登錄會員、繳交信用卡費、看醫生、餐廳訂位以及使用臉書打卡等，人們容易使用智慧型手機或是平板電腦等攜帶式行動裝置完成，但因使用上的方便，人們也越來越容易在網路或是雲端資料庫中，留下許多“痕跡”，其又可稱為資料。雖然私人資料係有法律規範所保障，但許多消費者的使用痕跡容易被拿來運用及分析，分析者則容易由資料中找出發展的趨勢及消費者的喜好程度改變，以利快速面對日新月異的改變。

【0003】 近年來，大數據（Big data或Megadata）又可稱巨量資料、海量資料或大資料，其皆係為所涉及的資料量規模巨大到無法透過人工在合理時間內達到擷取、管理、處理及整理成為一般人所能解讀之資訊。在總資料量相同的情況下，與個別分析獨立的小型資料相比，將各個小型資料合併後進行分析則可得出許多額外的資訊和資料關聯性，可以用來找出商業趨勢、判定研究品質、避免疾病擴散、打擊犯罪或測定即時交通路況等資料，但每一次的資料整理皆需耗費許多的人工或是時間，以執行一個新的資料分析，小型資料還容易計算，但在面對現今巨量資料時，舊有的統計方式會來不及面對“改變”。

【0004】 有鑑於此，本發明係在針對昔知之缺失，提出一種藉由已知資料的資料特徵以使新的資料學習的遷移學習演算方法，以同時改善上述的該些問

題。

【發明內容】

【0005】 本發明的主要目的係在提供一種遷移學習演算方法，藉由模糊演算法以及已知的資料所計算出的資料特徵，以讓新的或未知的資料進行資料分析的學習，當面臨巨量資料時，可以更快速的找出新的或未知的巨量資料的特徵，並且藉由已知資料的經驗更提高了資料分析的準確性，以讓分析資料的使用者更快速的掌握所分析之未知的資料特徵。

【0006】 爲了達到上述的目的，本發明提供一種遷移學習演算方法，其係包含下列步驟，先取得一筆已知資料並且隨機產生一第一初始化參數，並藉由一模糊演算法計算這筆已知資料，以獲得一第一資料特徵；將一筆未知資料遷移至模糊演算法中，並以第一資料特徵作爲此筆未知資料的初始化參數計算出一資料結果；及判斷資料結果是否符合模糊演算法的收斂條件，若是，此筆未知資料可以快速學習已知資料的資料結果計算以直接產生一第二資料特徵；若否，則將資料結果重新加入至此筆未知資料的初始化參數中，並以模糊演算法計算及判斷，重複計算及判斷以計算出一更新資料結果以符合模糊演算法的收斂條件。

【0007】 底下藉由具體實施例配合所附的圖式詳加說明，當更容易瞭解本發明之目的、技術內容、特點及其所達成之功效。

【圖式簡單說明】

【0008】

第一圖爲本發明的架構示意圖。

第二圖爲本發明遷移學習演算方法的步驟流程圖。

第三圖爲第二圖中新增一步驟的步驟流程圖。

【實施方式】

【0009】 模糊演算法又可稱為模糊(Fuzzy)理論，其係以模糊集合(Fuzzy Set)為基礎以接受模糊性現象存在的事實，而處理概念模糊不確定的事物為研究方向，並且積極的將量化成電腦的程式及軟體可以處理的訊息資料，Fuzzy理論的應用比較偏重於人類的經驗及對問題特性的掌握程度，以使分析者可以快速掌控資料特徵。

【0010】 本發明係提供一種遷移學習演算方法，在說明方法之前，先說明本發明的系統架構。首先，請先參照本發明第一圖所示，一已知資料10藉由演算單元12計算，並計算出已知資料10的第一資料特徵，其中演算單元12係可使用電腦程式，其係為可使用模糊演算法演算的程式，在本實施例中演算單元12係可為電腦，電腦程式係為矩陣實驗室數學軟體(Matrix Laboratory, MATLAB)，而模糊演算法係為模糊C均值(Fuzzy C-means, FCM)演算法；一未知資料14於演算單元12中，使用已知資料10的第一資料特徵作為初始化參數計算，以產生第二資料特徵，且再將第二資料特徵送至輸出單元16，以供分析者取得計算後的未知資料14的第二資料特徵。

【0011】 說明完本發明的系統架構之後，接續詳述本發明遷移學習演算方法的步驟流程，請參照本發明第二圖所示，並請同時再參照第一圖。首先，如步驟S10所示，先取得一筆已知資料10及藉由演算單元12隨機產生一第一初始化參數，並藉由模糊演算法計算已知資料以獲得第一資料特徵，本實施例中模糊演算法係為模糊C均值演算法。如步驟S12所示，將一筆未知資料14遷移至模糊演算法中，並以第一資料特徵作為未知資料14的初始化參數，其中係根據第一資料特徵的一平均值及一變異數作為未知資料14的初始化參數，並以計算出一資料結果。如步驟S14所示，判斷資料結果是否符合模糊演算法的收斂條件，若是則跳下一步驟。如步驟S16所示，未知資料14可以快速學習已知資料10的資料結果，並計算已直接產生一第二資料特徵，以達成計算後的未知資料14的輸出

結果，並可由輸出單元16供分析者取得計算後的未知資料14的第二資料特徵。若否則跳到步驟S18，如步驟S18所示，則將資料結果重新加入至未知資料14的初始化參數中，並以演算單元12使用模糊演算法計算及判斷，重複計算及判斷以計算出一更新資料結果，並可以符合模糊演算法的收斂條件。

【0012】 承接上段，其中模糊C均值演算法中，參數J代表目標函數且係在群組中整個平方誤差的和(overall within-group sum)，參數C代表資料的平均值，參數i及j代表資料中的各個點，參數n代表資料的群，參數m代表權種係數，參數u代表點與群的相關程度，參數x代表輸入資料，參數v代表資料的中心點，參數D代表每一點與中心點的距離，模糊C均值演算法係經由公式(1)及公式(2)計算出：

$$J = \sum_{i=1}^C \sum_{j=1}^n u_{ij}^m D_{ij} \quad (1)$$

$$D_{ij} = \|x_j - v_i\|^2 \quad (2)$$

再者，步驟S14中的收斂條件中係由公式(3)所計算出：

$$\sum_{i=1}^C u_{ij} = 1, \forall j \in [1, n] \quad (3)$$

【0013】 藉由本發明的遷移學習演算方法可以使資料與資料間作更快速學習，經由已演算過的已知資料的技術特徵作為未知資料的學習條件，可以加速未知資料的演算速度，並且提高演算後的輸出資料的準確度，尤其面對巨量資料時，使用者能更快速及準確的掌握資料的特徵。

【0014】 再者，請參照本發明第三圖所示，第三圖中的步驟S10~S18相同於第二圖的步驟S10~S18，在此則不贅述，上述第二圖中的未知資料及已知資料係為相關的資料，倘若未知資料與已知資料係完全不相關時，則無法直接計算出未知資料的資料結果。如步驟S20所示，資料結果經重複計算及判斷後，仍然無法符合模糊演算法的的收斂條件時，則可將未知資料及藉由演算單元隨機產

生一第二初始化參數，並藉由演算單元以模糊演算法計算未知資料，以獲得第三資料特徵。

【0015】 以上所述之實施例僅係為說明本發明之技術思想及特點，其目的在使熟習此項技藝之人士能夠瞭解本發明之內容並據以實施，當不能以之限定本發明之專利範圍，即大凡依本發明所揭示之精神所作之均等變化或修飾，仍應涵蓋在本發明之專利範圍。

【符號說明】

【0016】

10 已知資料

12 演算單元

14 未知資料

16 輸出單元

【發明申請專利範圍】

【第1項】 一種遷移學習演算方法，包括下列步驟：

取得一筆已知資料及隨機產生一第一初始化參數，並藉由一模糊演算法計算該筆已知資料，以獲得一第一資料特徵；

將一筆未知資料遷移至該模糊演算法中，並以該第一資料特徵作為該筆未知資料的初始化參數，以計算出一資料結果；及

判斷該資料結果是否符合該模糊演算法的收斂條件：

若是，該筆未知資料可以快速學習該筆已知資料的該資料結果，並計算以直接產生一第二資料特徵；及

若否，則將該資料結果重新加入至該筆未知資料的該初始化參數中，並以該模糊演算法計算及判斷，重複計算及判斷以計算出一更新資料結果以符合該模糊演算法的收斂條件。

【第2項】 如請求項1所述之遷移學習演算方法，更包括一步驟，其係在重複計算及判斷後無法符合該模糊演算法的收斂條件時，則將該筆未知資料及隨機產生一第二初始化參數，並藉由該模糊演算法計算該筆未知資料，以獲得一第三資料特徵。

【第3項】 如請求項2所述之遷移學習演算方法，其中該模糊演算法係為模糊C均值（Fuzzy C-means，FCM）演算法。

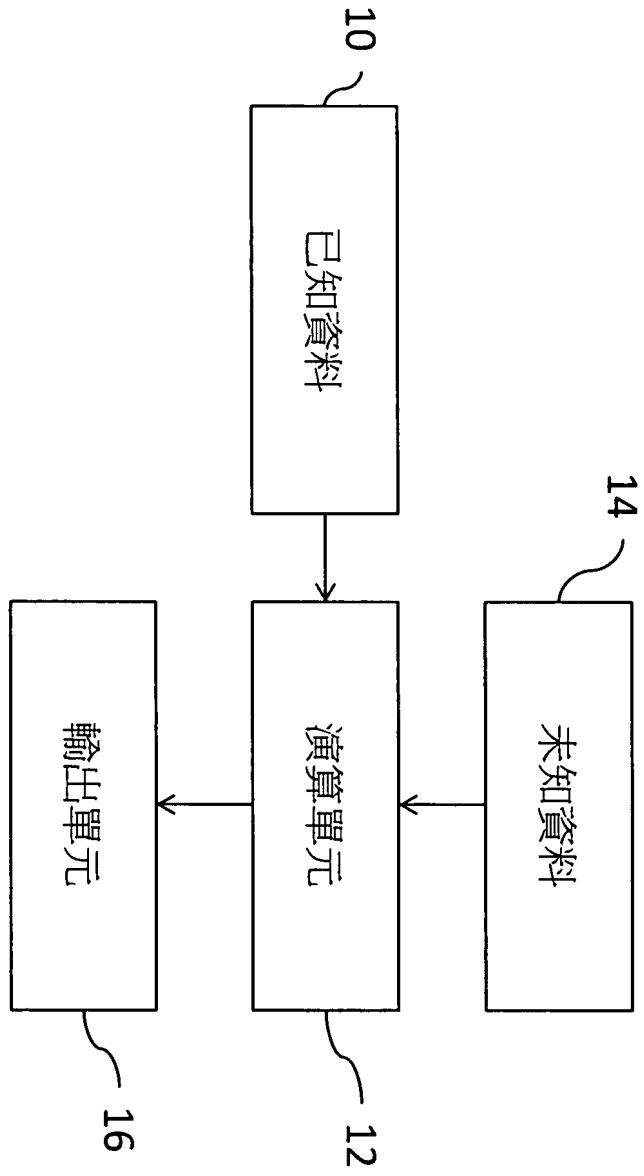
【第4項】 如請求項3所述之遷移學習演算方法，其中該模糊C均值演算法的計算公式係為 $J = \sum_{i=1}^C \sum_{j=1}^n u_{ij}^m D_{ij}$ ， $D_{ij} = \|x_j - v_i\|^2$ ，其中上述公式之各參數分別為：J代表目標函數，C代表平均值，i及j代表各個點，n代表群，m代表權種係數，u代表點與群的相關程度，x代表輸入資料，v代表中心點，D代表每一點與中心點的距離。

【第5項】 如請求項1所述之遷移學習演算方法，其中根據該第一資料特徵的一

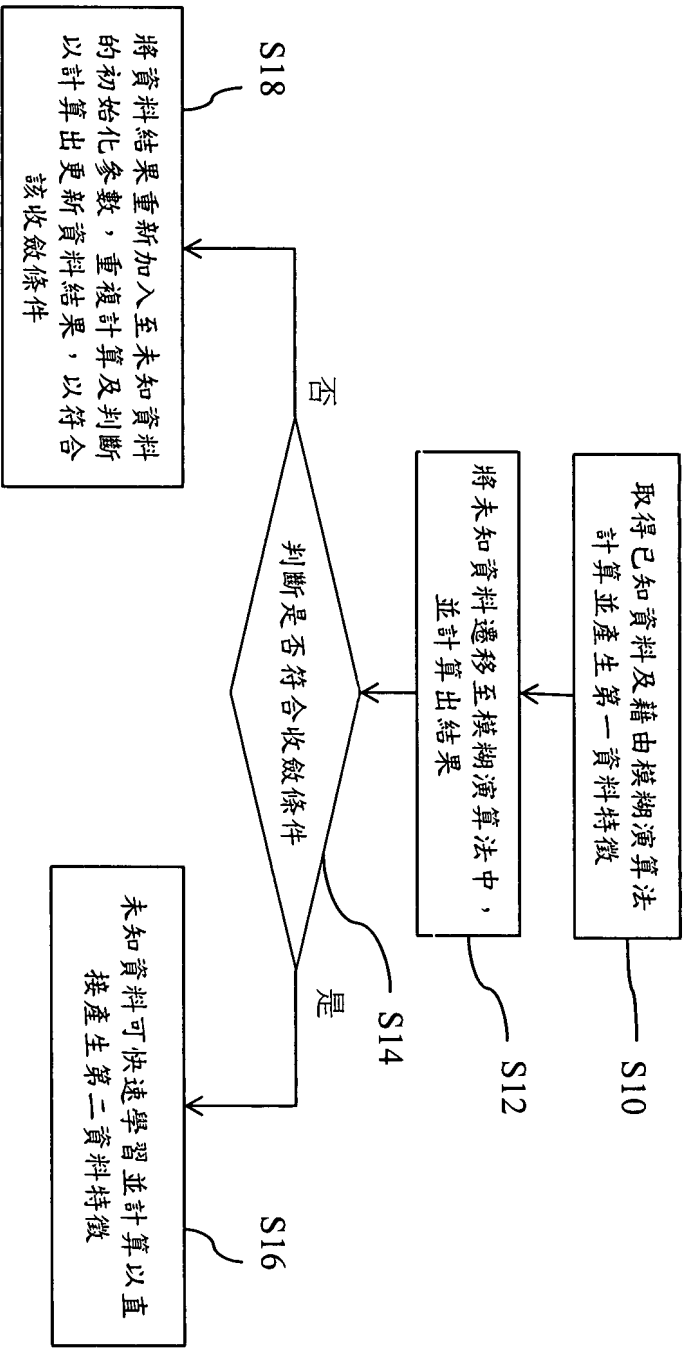
平均值及一變異數以作為該筆未知資料的該初始化參數。

- 【第6項】 如請求項1所述之遷移學習演算方法，其中該收斂條件係為極值條件，其係為 $\sum_{i=1}^c u_{ij} = 1, \forall j \in [1, n]$ ，其中上述公式之各參數分別為：C代表平均值，i及j代表各個點，n代表群。
- 【第7項】 如請求項1所述之遷移學習演算方法，其中該資料結果符合該模糊演算法的收斂條件時，該筆未知資料係與該筆已知資料為相關的資料。
- 【第8項】 如請求項2所述之遷移學習演算方法，其中在重複計算及判斷後無法符合該模糊演算法的收斂條件時，該筆未知資料係與該筆已知資料為不相關的資料。
- 【第9項】 如請求項2所述之遷移學習演算方法，其中該第一初始化參數係藉由一電腦程式以隨機產生。
- 【第10項】 如請求項9所述之遷移學習演算方法，其中該第二初始化參數係藉由該電腦程式以隨機產生。
- 【第11項】 如請求項10所述之遷移學習演算方法，其中該電腦程式可以進行該模糊演算法。
- 【第12項】 如請求項11所述之遷移學習演算方法，其中該電腦程式係為矩陣實驗室數學軟體（Matrix Laboratory，MATLAB）。

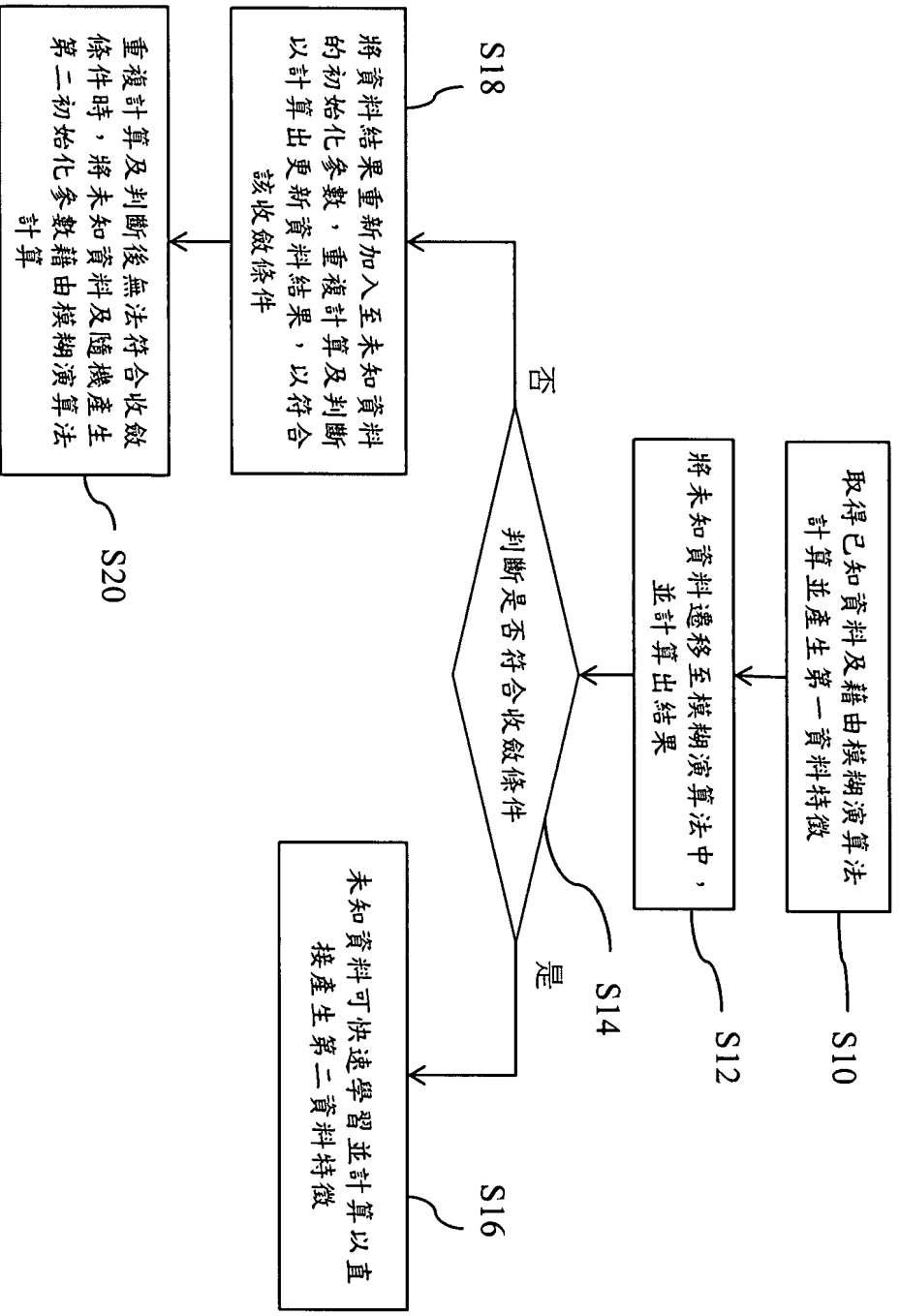
【發明圖式】



第一圖



第二圖



第三圖