

219993

公告本

申請日期	8/3/21
案號	219993
類別	G10L 2/08

A4
C4

(以上各欄由本局填註)

發明專利說明書

(請先閱讀背面之注意事項再填寫本頁各欄)

一、發明名稱	中文	獨立式及時語音辨識系統
	英文	
二、發明人	姓名	謝朝凱
	籍貫 (國籍)	中華民國
	住、居所	新竹縣芎林鄉上山村6鄰58號
三、申請人	姓名 (名稱)	財團法人工業技術研究院
	籍貫 (國籍)	中華民國
	住、居所 (事務所)	新竹縣竹東鎮中興路4段195號
	代表人 姓名	林垂宙

經濟部中央標準局員工消費合作社印製

裝
訂
線

五、發明說明(2)

5-3 圖式說明：

- 圖一、本發明獨立式及時語音辨識系統方塊圖
 圖二、本發明獨立式及時語音辨識系統軟體流程圖
 圖三、本發明之特徵點萃取模式圖
 圖四、本發明之語音端點偵測模式圖
 圖五、傳統之動態規畫比對法
 圖六、本發明所使用之對稱法
 圖七、本發明之動態規畫比對法
 圖八、本發明之一維循環式緩衝器
 圖九、本發明之最佳實施例

圖號說明：

- | | |
|----------------|------------------|
| 101 麥克風 | 102 運算放大器 |
| 103 類比對數位轉換器 | 104 數位信號處理器(DSP) |
| 105 參考樣本存放器 | 106 控制電路 |
| 107 辨識器 | 201 前級處理步驟 |
| 202 特徵點萃取步驟 | 203 語音端點偵測步驟 |
| 204 辨認步驟 | 205 有效參考樣本步驟 |
| 206 辨認比對步驟 | 901. 16位元位移暫存器 |
| 902 16位元緩衝器 | 903 時脈產生器 |
| 904 串聯轉並聯旗標產生器 | 905 辨識器 |
| 906 16位元緩衝器 | 907 參考樣本存放器 |
| 908 解碼器 | 909 解碼器 |
| 910 16位元位址計數器 | 911 16位元緩衝器 |

(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)

5-4 發明詳細說明：

本發明藉著修正過的線性預測編碼(LPC)及修正過的動態規劃比對(DTW)，藉著一個簡易的數位信號處理器(例如TMS32010)，即可完成一個不須附屬在另一個處理機上的獨立式即時語音辨識系統。

由於此系統乃屬於特定語者的辨識系統，所以必須經過二個階段(一)訓練過程(二)辨識過程，才能完成語音及時辨認的工作。

為了詳細說明此獨立式即時語音辨識系統，以下分為硬體架構(圖一)及軟體流程(圖二)來說明。

在硬體方面如圖一所示為本發明之方塊圖，麥克風101為拾取語音並將語音轉為類比式電平訊號，運算放大器102將由麥克風所傳來微弱的電平訊號放大；類比對數位轉換器(A/D converter) 103將由運算放大器102所傳來的類比信號轉換為數位語音信號以供給數位信號處理器(DSP)104處理之用；參考樣本存放器105用來儲存在訓練過程中所取得的參考樣本(reference Pattern)以供未來在辨認過程中使用，控制電路106為控制類比對數位轉換器103及數位信號處理器104之間數位信號的串聯一並聯轉換及產生時脈電路(Timing generator)，辨識器107為用來存放有關

(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)

語音訊號特徵點的萃取(feature extraction)、端點偵測(endpoint detection)、修正過的動態規劃比對(DTW)等資料以提供數位信號處理器(DSP)104，在訓練過程及辨識過程之用，最後的辨認結果由數位信號處理器104輸出。

在軟體方面如圖二所示，數位語音信號先經過前級處理步驟201的過程，也就是先經過一階濾波(first order filter) $(1-0.937Z^{-1})$ ，將數位語音信號中的高頻語音信號部份加強，以避免高頻語音信號衰減，之後在特徵點之萃取步驟202，將已加強高頻語音信號之數位語音信號每隔30毫秒(ms)並重疊10毫秒為一音框(frame)如圖三所示，並經過取窗函數的近似漢明取窗(Hamming Window)式子如下

$$h(n)=0.5(1-\cos(2\pi n/(N-1))) \quad 0 \leq n \leq N-1$$

$$= 0 \quad \text{otherwise}$$

再利用杜賓(Durbin)化簡法，以求取10階固定點(fixed point)的線性預測係數(LPC)為參數，以供辨認之用。

在語音端點偵測步驟203之偵測中，利用上述在求10階固定點的線性預測係數之過程中，所求得的能量參數，以尋找出語音部份並去除雜訊(noise)的部份。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(5)

在辨認步驟204中，可分為訓練過程及辨認過程，如為訓練過程則將上述10階固定點線性預測係數存入參考樣本存放器105中做為參考樣本，即執行有效參考樣本(Referencre Pattern)步驟205；如為辨認過程則執行辨認比對步驟206，以對稱法(Symmetric form)修正過的動態規劃比對(DTW)以求取辨認結果，其辨認方式為將所欲尋找最短距離以一維(one dimention)循環式(Circular)的緩衝器來製作，而最短距離即為辨認結果；而此緩衝器所須的記憶體大小僅約10到13個左右即可，所以比傳統上的技術可以大量減少記憶體的使用；在求最短距離的比對，則採下列公式：

$$d(i, j) = | a_i - b_j |$$

a_i : 參考樣本之LPC參數
 b_j : 欲辨認樣本之LPC參數

在語音端點偵測步驟(end point detecetion)203中，利用能量參數，以尋找出語音部份並去除雜訊部份(noise)如圖四所示，D為波峰的寬度亦為單音的長度，BTW為兩波峰之間的寬度亦為兩單音之間的長度，其方法如下：

步驟1：針對每一音框(frame)，求取能量參數E；

$$E = \sum_{i=1}^{240} [S(i)]^2, \quad S(i) \text{ 表示每一筆A/D的數位}$$

語音信號

步驟2：設定總長L為 ϕ ，抓取一能量參數E；

步驟3：假如E小於臨界(Threshold)值，則該音框(frame)為雜訊，抓取下一音框(frame)的能量參數直到E大於臨界(Threshold)值；

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(6)

步驟4：設定旗標(flag)=0 (表示單音)；

步驟5：設定D=0；

步驟6：假如D大於臨界值，D長度增加1，並抓取下一音框，直到D小於臨界值；

步驟7：L=L+D

步驟8：Case 1：假如旗標(flag)等於0

假如D小於8，則視為雜訊，將L清除為0，並回到步驟1

假如D大於等於8，則設BTW為0，旗標為1，執行步驟9

Case 2：假如旗標(flag)等於1

假如D小於8，則BTW=BTW+D

假如D大於，等於8，則設定BTW為0；

步驟9：假如[(E小於臨界值)並且(BTW小於等於16)]則BTW=BTW+1 並抓取下一音框，回到步驟9；

步驟10：假如BTW小於16則L=L+BTW，回到步驟5；

步驟11：L=L-BTW，清除(clear)BTW，輸出語音總長L；

步驟12：結束語音端點偵測(end point detection)。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(7)

傳統動態規劃比對(DTW)方法如圖五所示，本發明係採用以如圖六所示之對稱法(Symmetric form)修改過之圖七動態規劃比對法，來做辨認比對步驟204，不過由於比對路徑函數(wrapping function)一般皆用二維方式(two dimension)來劃作，從圖五我們知道其所需儲存量(wrapping space)是 $i \times j$ ，當 i, j 甚大時，將會浪費許多記憶體儲存，為了節省記憶體我們修改其二維函數，成為一維函數，其儲存量最多只須 $2W+1$ (W 為adjust window size)， W 值在本發明之實驗值為6；此外為了避免尋找路徑過長，使得此一維函數會發生溢流(over flow)情形，我們用一如圖8所示之循環緩衝器(circular buffer)來製作。為了說明以上之以對稱法修正過的動態規劃比對法，我們先定義以下變數：

i ：測試樣本的長度

j ：參考樣本的長度

st ：對 i 而言搜尋範圍的起始點

ed ：對 i 而言搜尋範圍的終點

bef ：搜尋長度($ie, ed-st+1$)

Sti ：前一個搜尋的起始點X座標值

Stj ：前一個搜尋的起始點Y座標值

Ptr ：搜尋長度或搜尋長度+1

W ：窗大小(window size)

Ptg ：循環緩衝器(circular buffer)的指標(pointer)

DTW ：總長度(Total distance)

a_i ：測試樣本之LPC參數

b_j ：參考樣本之LPC參數

g ：循環緩衝器大小(circular buffer size) $=2 \times window+1$

$d(i, j)$ ：表示兩點間距離

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(8)

步驟1：設定*i, j, sti, stj, Ptg*為1
設定 $bef = 1 + window$

步驟2： $i = i + 1$

步驟3：If ($i > j + window$)

then 計算 $st = j - window$;

$ed = j + window$;

If ($ST \leq 0$) $ST = 1$;

If ($ed >$ 測試樣本總長) $ed =$ 測試樣本總長

If (現在搜尋的起始座標各減1等於 ($Sti,$
 Stj) 座標)

then $ptr = bef$;

else $ptr = bef + 1$;

If ($j >$ 參考樣本總長)

then $DTW = sum / c$ 測試樣本總長度 + 參考樣
本總長度)

$Ssop$;

else $i = j - window$

回到步驟3

else

If ($i \leq 0$) 回到步驟2

If ($i >$ 測試樣本總長) 回到步驟(step)2

$pth = ptg + 1$

$$d(i, j) = \sum_{m=1}^{10} |a_{mi} - b_{mj}|$$

If ($i \geq 1$ and $j - 1 \geq 1$ and $|i - j + 1| <$
 $= window$)

then 路徑 $l = g[ptg - ptr + 1] + d(i, j)$;

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(9)

```

If(i-1>=1 and j-i>=1)
then路徑2=g[ptg-ptr]+2*d(i,j);
If(i-1>=1 and j>=1 and |i-j-1|
<=window)
then路徑3 =g[ptg-1]+d(i,j);
g[Ptg]=minimum(路徑1, 路徑2, 路徑3)
回到步驟2。

```

5-5 最佳實施例：

如圖 9 所示為本發明之最佳實施例，類比語音信號經由麥克風 101 送到運算放大器 102，將類比語音信號放大到 +5V—-5V 之間，再經由類比轉數位轉換器 (A/D converter) 103，將類比語音信號轉成串列式 (serial) 的 8 位元 log-PCM 的數位信號 (PCMOUT)，再將此數位信號 (PCM OUT) 送到一 16 位元的位移暫有器 (shift register) 901 內，如此則可將串列式 8 位元 log-PCM 的數位信號，轉成並列式 (parallel) 以便數位信號處理器 (DSP) 104 處理。

在取樣頻率方向，則由時脈產生器 (clock generator) 903 產生時脈以供應：(一) 主要時鐘脈波 (master clock) CLKIN 20 MHz 給數位信號處理器 (DSP) 104 用。(二) 2MHz 之 CLKX/CLKR 及 8KHz 之 FSX/FSR 給數位轉類比轉換器 103 用 (三) 另產生一脈波 QD 以傳給串聯轉並聯旗標產生器 (Serial-to-Parallel flag generator) 904 之用。當 16 位元位移暫存器 901 接收到上述 2 筆 8 位元 log-PCM 的數位信號後由串列轉並列旗標產生器 904 所產生之 BIO 信號會通知數位信號處理器 (DSP) 104 以接收由 16 位元位移暫存器 901 及 16 位元緩衝器 902 所傳送之 16 位元並列式語音數位信號，接著，數位信號處理器 (DSP) 104 及辨識器 (recognizer, 4K×16ROM) 905 及 16 位元緩衝器 (16bit buffer) 906 將數位語音信號經過

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

五、發明說明(10)

一階濾波(forst order filter) (1-0.937Z)之前級處理201，接著再做特徵點之萃取202每隔30毫秒重疊(overlap) 10毫秒分析產生一組10階固定點線性預測係數為辨認參數，再進行語音端點偵測203，如為訓練過程則將上述之辨認參數放入參考樣本存放器(reference storge 32K×16 SRAM) 907，如為辨認過程，解碼器908及解碼器909兩組解碼器(decoder) 會產生驅動信號(drive signal)CP及load給16位元位址計數器(16 bit adres scounter) 910以產生參考樣本存放器(907) 所需的位址(address)，並將此位址內所取出之辨認參數經由16位元緩衝器(16 bit buffer)送出，並運用辨識(905)內的以對稱法修正過的動態規畫比對法，在數位信號處理器104中經由比對，以求取辨認結果的輸出。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

四、中文發明摘要(發明之名稱：

獨立式及時語音辨識系統)

摘要：

一種獨立式及時語音辨識系統，包含麥克風可將語音訊號轉為類比式電平訊號，並以放大器放大，再以類比轉數位轉換器將類比語音訊號轉為一串列式數位信號，並將此串列式數位信號轉為並列式數位信號，數位處理器則將並列式數位信號加以做前級處理，特徵點萃取及語音端點偵測並得到一多位階固定點線性預測係數，在訓練過程中此一多位階固定點線性預測係數將被儲存做為參考樣本，在辨識過程則此一多位階固定點線性預測係數將被以對稱法修正過的動態規畫比對法和參考樣本比對以得到一辨識結果的輸出。

英文發明摘要(發明之名稱：

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

附註：本案已向

國(地區)申請專利，申請日期：

案號：

六、申請專利範圍

8年 月 日 修正 補充

申請專利範圍：

1. 一種獨立式及時語音辨識系統，本系統包含：

麥克風裝置係將語音訊號轉為類比式電平訊號；

放大裝置將上述之類比式電平訊號之電壓放大以產生一放大類比式電平訊號；

類比轉數位轉換裝置將上述之放大類比式電平訊號轉換為一串列式數位信號；

控制裝置將上述之串列式數位信號轉為並列式數位信號及產生時脈信號；

辨識裝置儲存前級處理資料、特徵點萃取資料、端點偵測資料及以對稱法修正過的動態規畫比對資料；

參考樣本存放裝置儲存參考樣本；

數位信號處理裝置藉著上述之前級處理資料，上述之特徵點萃取資料及上述之端點偵測資料來對上述之並列式數位信號處理以產生一多位階固定點線性預測係數並去除上述語音訊號中之雜訊部份；

在訓練過程上述之多位階固定點線性預測係數被送至上述之參考樣本存放裝置以做為上述之參考樣本；

在辨識過程上述之多位階固定點線性預測係數被以上述之以對稱法修正過的動態規畫比對資料和上述之參考樣本在數位信號處理裝置中比對以得到一辨認結果並將此辨認結果輸出。

(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

2. 如申請專利範圍第 1 項之系統，其中上述之串列式數位信號為多位元 log-PCM 數位信號。

3. 如申請專利範圍第 1 項之系統，其中上述之控制裝置更包含：

移位暫存裝置將上述之串列式數位信號做移位暫存後送至一第一緩衝裝置，並由此第一緩衝裝置輸出上述之並列式數位信號；

時脈產生裝置產生第一時脈給上述之數位信號處理裝置使用，產生第二時脈給上述之類比轉數位轉換裝置使用，產生第三時脈給上述之串列轉並列旗標產生裝置使用。

4. 如申請專利範圍第 3 項之系統其中上述之辨識裝置更包含：

唯讀記憶體裝置，儲有上述之前級處理資料，上述之特徵點萃取資料，上述之端點偵測資料及上述之以對稱法修正過的動態規畫比對資料；

第二緩衝裝置，用以儲存上述之唯讀記憶裝置所輸出之資料，並將此輸出資料送入上述之數位信號處理裝置。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

封

六、申請專利範圍

5. 如申請專利範圍第 3 項之系統，其中上述之參考樣本存放裝置更包含：

解碼裝置產生一驅動信號給一位址計數器裝置以產生上述之參考樣本存放裝置之位址；

隨機取樣記憶體裝置，用以儲存參考樣本；

第三緩衝裝置暫存由上述之參考樣本存放器所輸出之上述之參考樣本，並將參考樣本輸出給上述之數位信號處理裝置。

6. 一種獨立式及時語音辨識方法，係將數位語音信號在訓練過程中轉為參考樣本，在辨識過程中輸出辨識結果，本方法包含下列步驟：

(a) 前級處理將上述之數位語音信號一階濾波以加強數位語音信號的高頻部份；

(b) 特徵點萃取將經前級處理之數位語音信號取音框並經過取窗函數的近似漢明取窗再用杜賓法化簡以得到一位階固定點線性預測係數；

(c) 語音端點偵測利用能量參數以尋找出語音部份並去除雜訊部份；

(d) 當訓練過程，儲存上述之多位階固定點線性預測係數做為上述之參考樣本，並回到(a)步驟；

(e) 當辨識過程，以對稱法修改過的動態規畫比對法將上述之多位階固定點線性預測係數和上述之參考樣本比對並輸出上述之辨識結果。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

219993

A7
B7
C7
D7

六、申請專利範圍

7. 如申請專利範圍第 6 項之方法，其中上述之取音框為每隔 30 毫秒並重疊 10 毫秒為一音框。
8. 如申請專利範圍第 6 項之方法，其中上述之多位階固定點線性預測係數為 10 位階固點線性預測係數。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

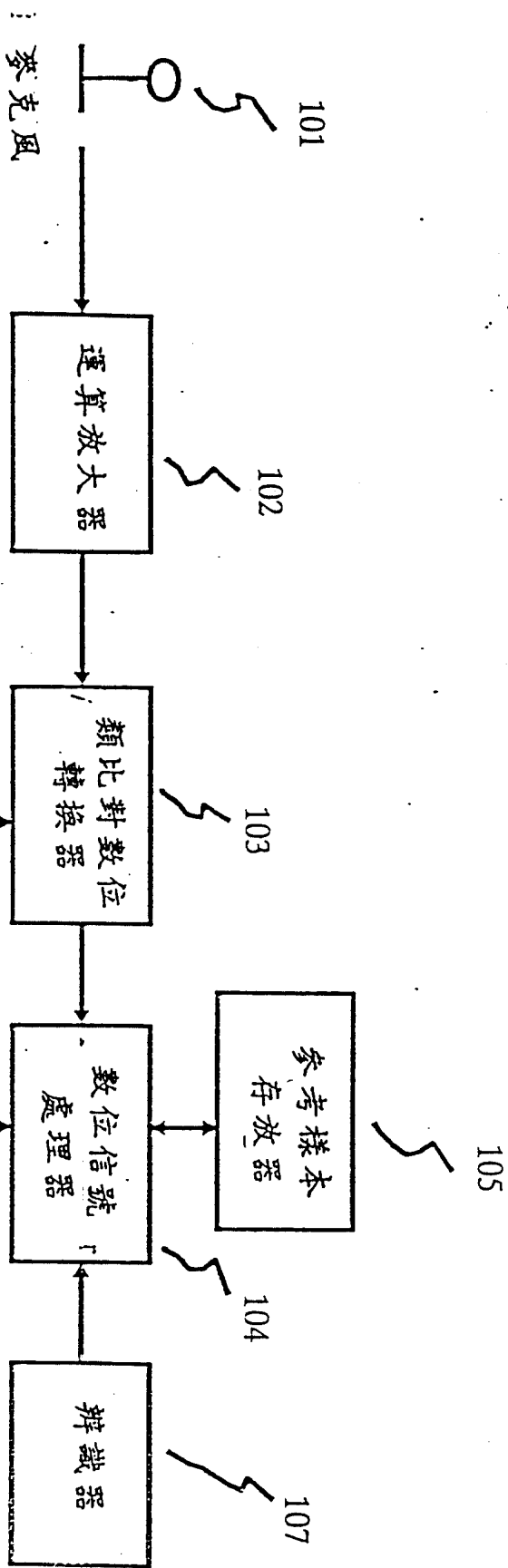
修正
補充
年 3 月 1 日

A8
B8
C8
D8

圖式

219993

圖



(請先閱讀背面之注意事項再行繪製)

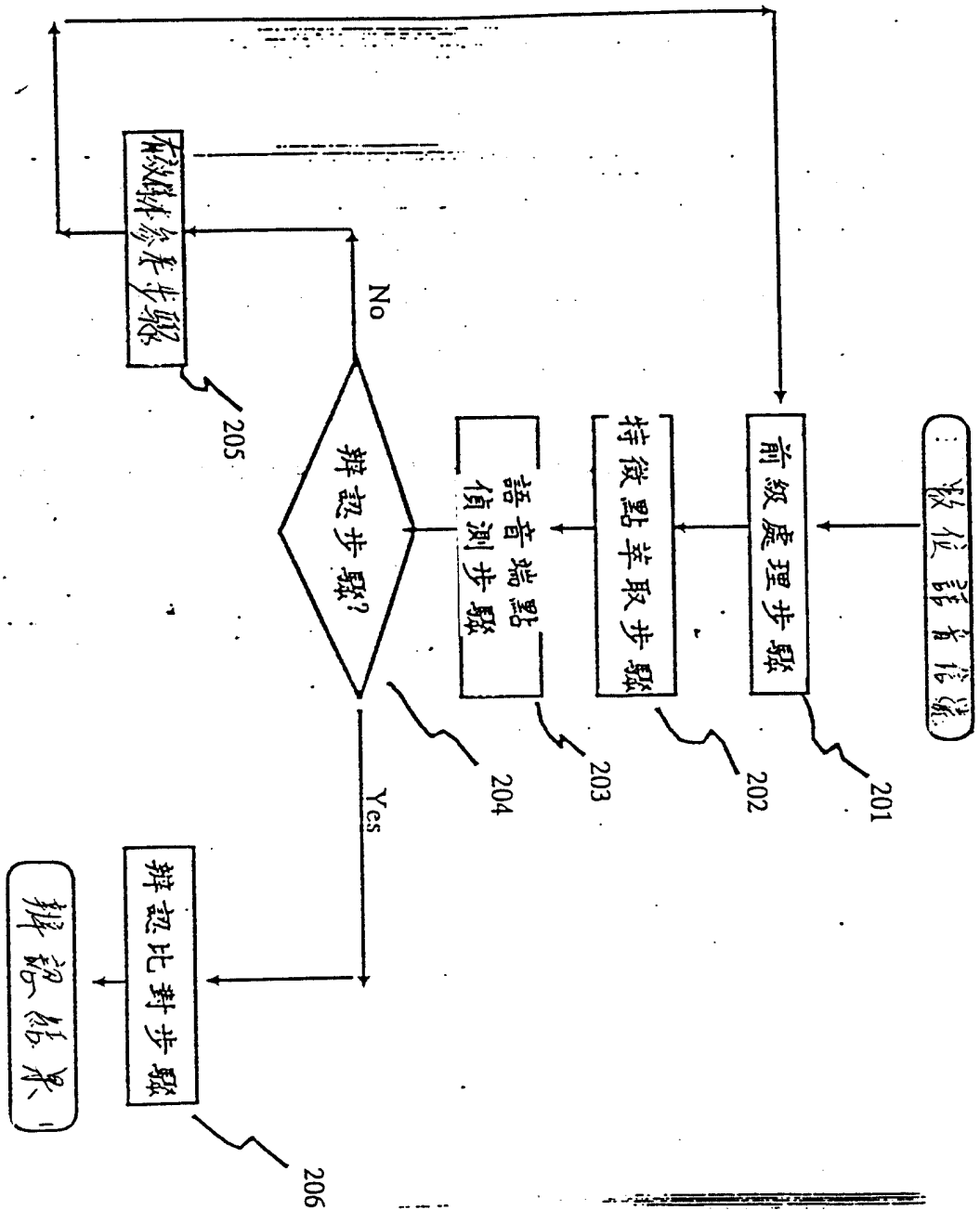
裝

訂

線

(請先閱讀背面之注意事項再行繪製)

圖式

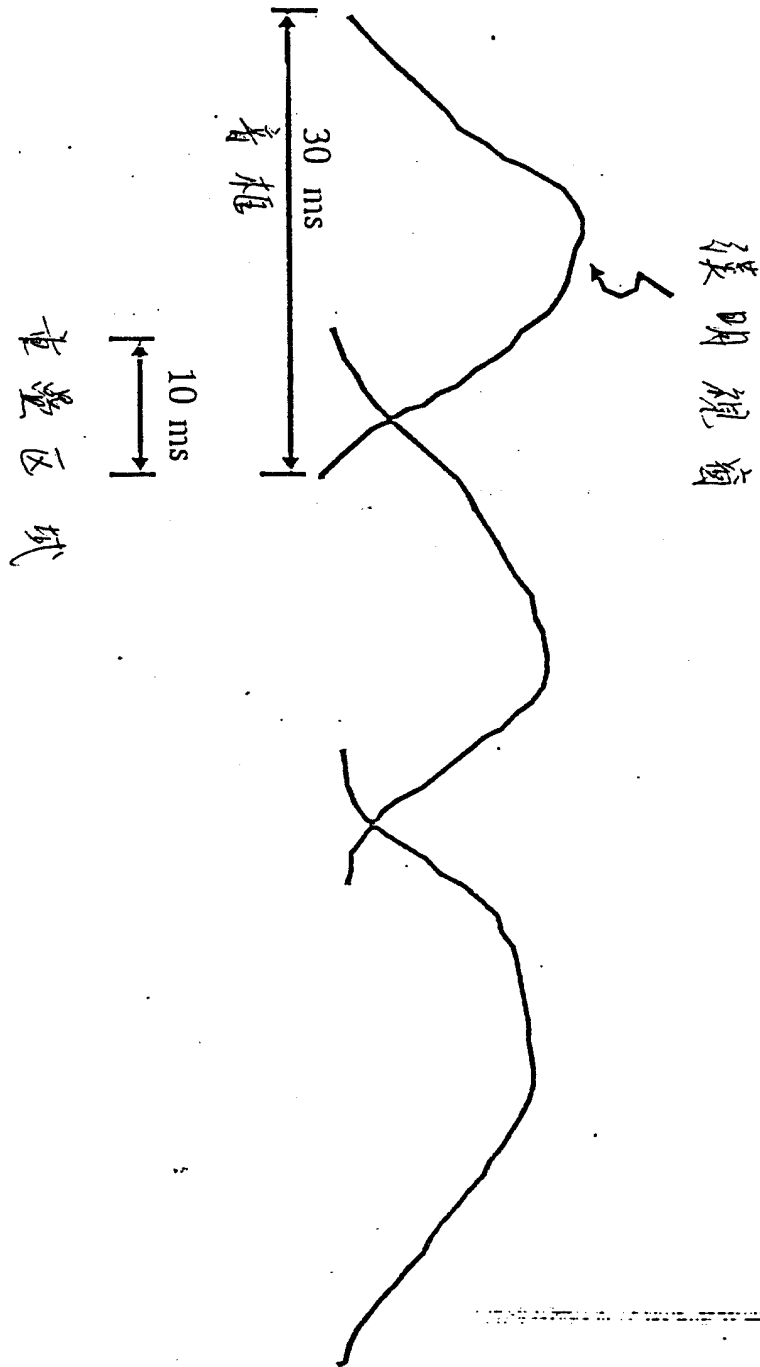


圖二

(請先閱讀背面之注意事項再行繪製)

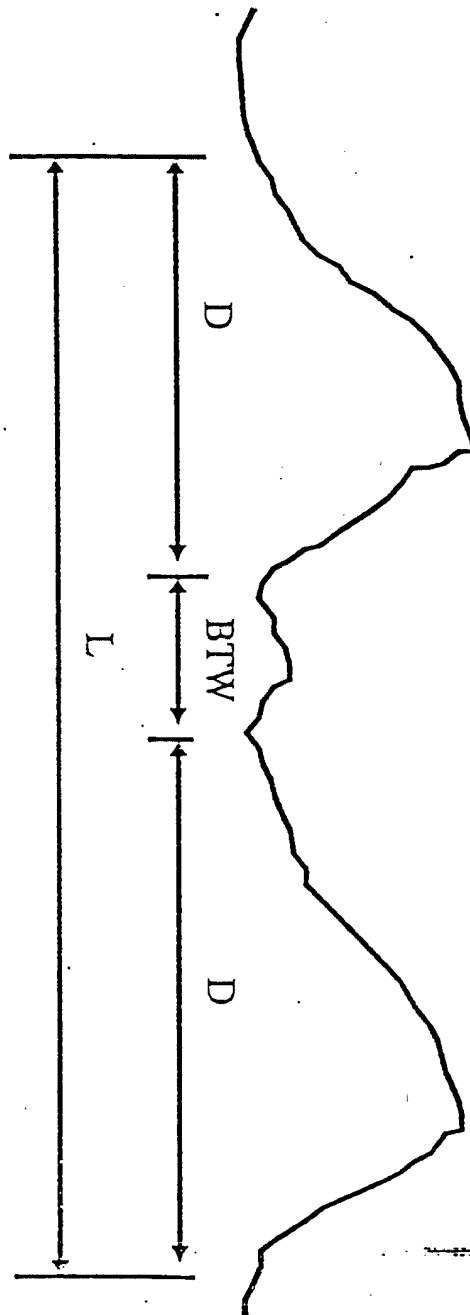
裝 訂 線

圖式



圖三

圖式



圖式

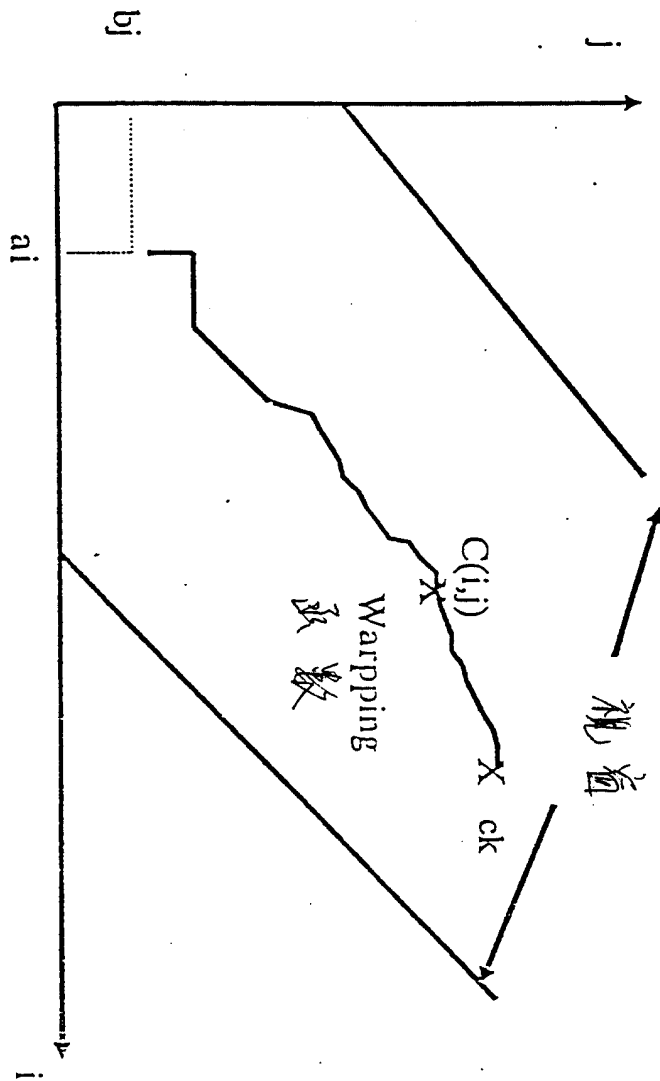
(請先閱讀背面之注意事項再行繪製)

裝

訂

線

圖式



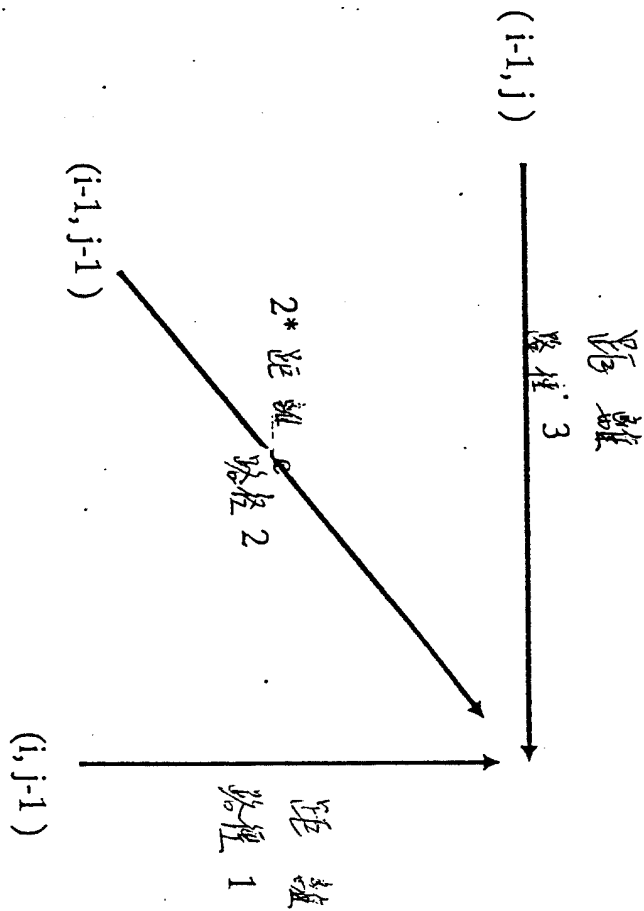
圖五

(請先閱讀背面之注意事項再行繪製)

裝 訂 線

(請先閱讀背面之注意事項再行繪製)

圖式

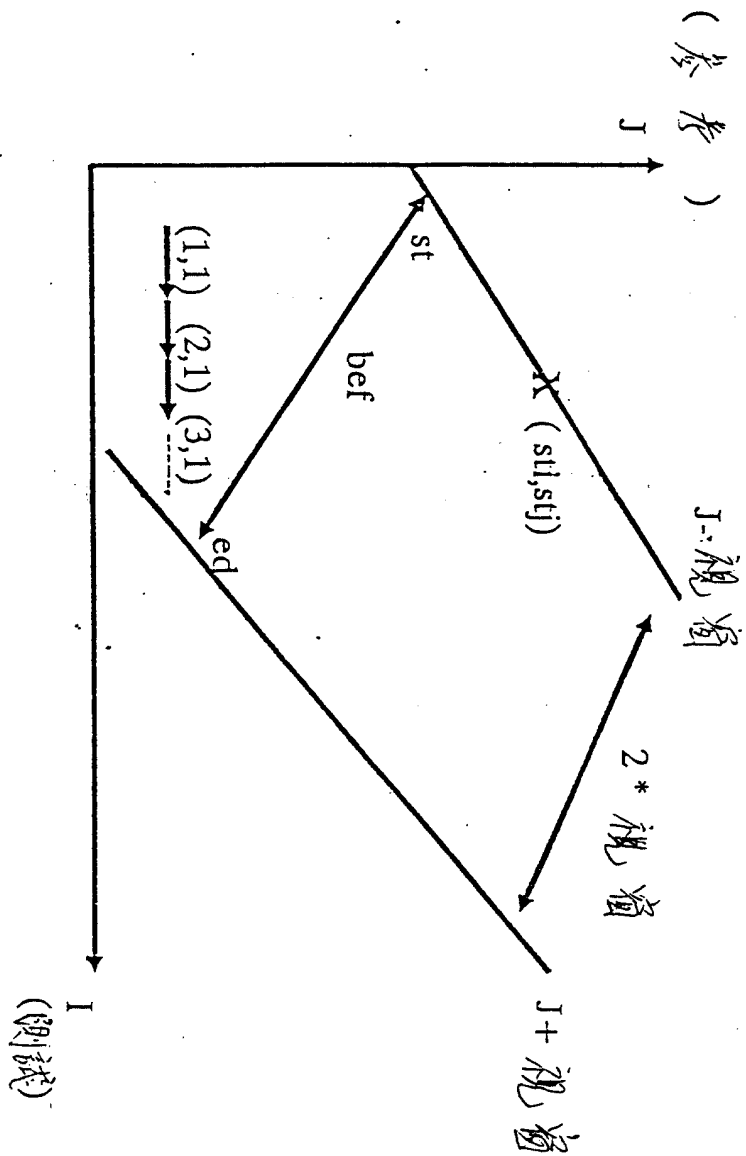


圖六

(請先閱讀背面之注意事項再行繪製)

裝 訂 線

圖式



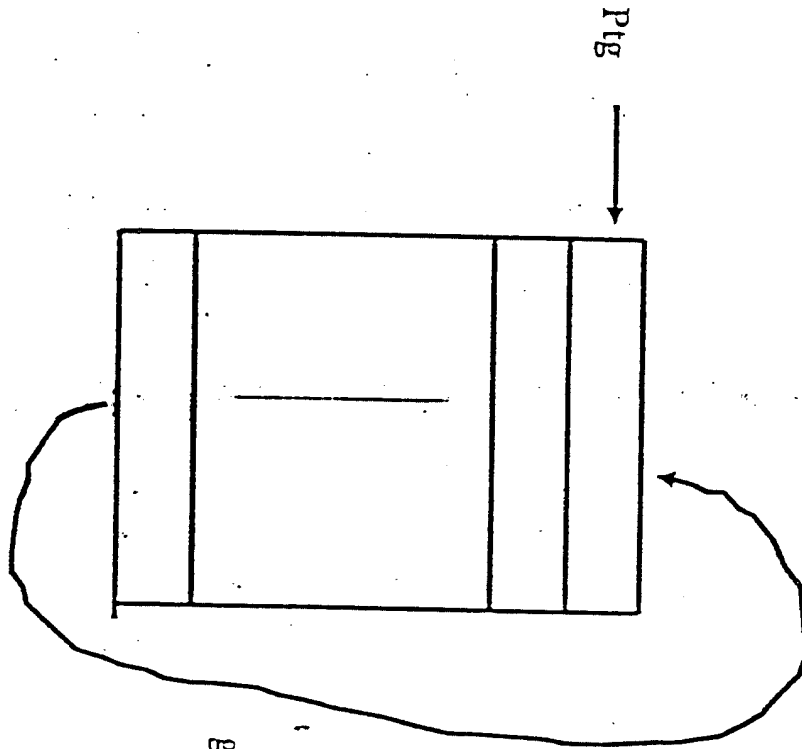
圖七

219993

A8
B8
C8
D8

圖式

圖八



螺旋式裝訂
B10: 2*視窗 1

(請先閱讀背面之注意事項再行繪製)

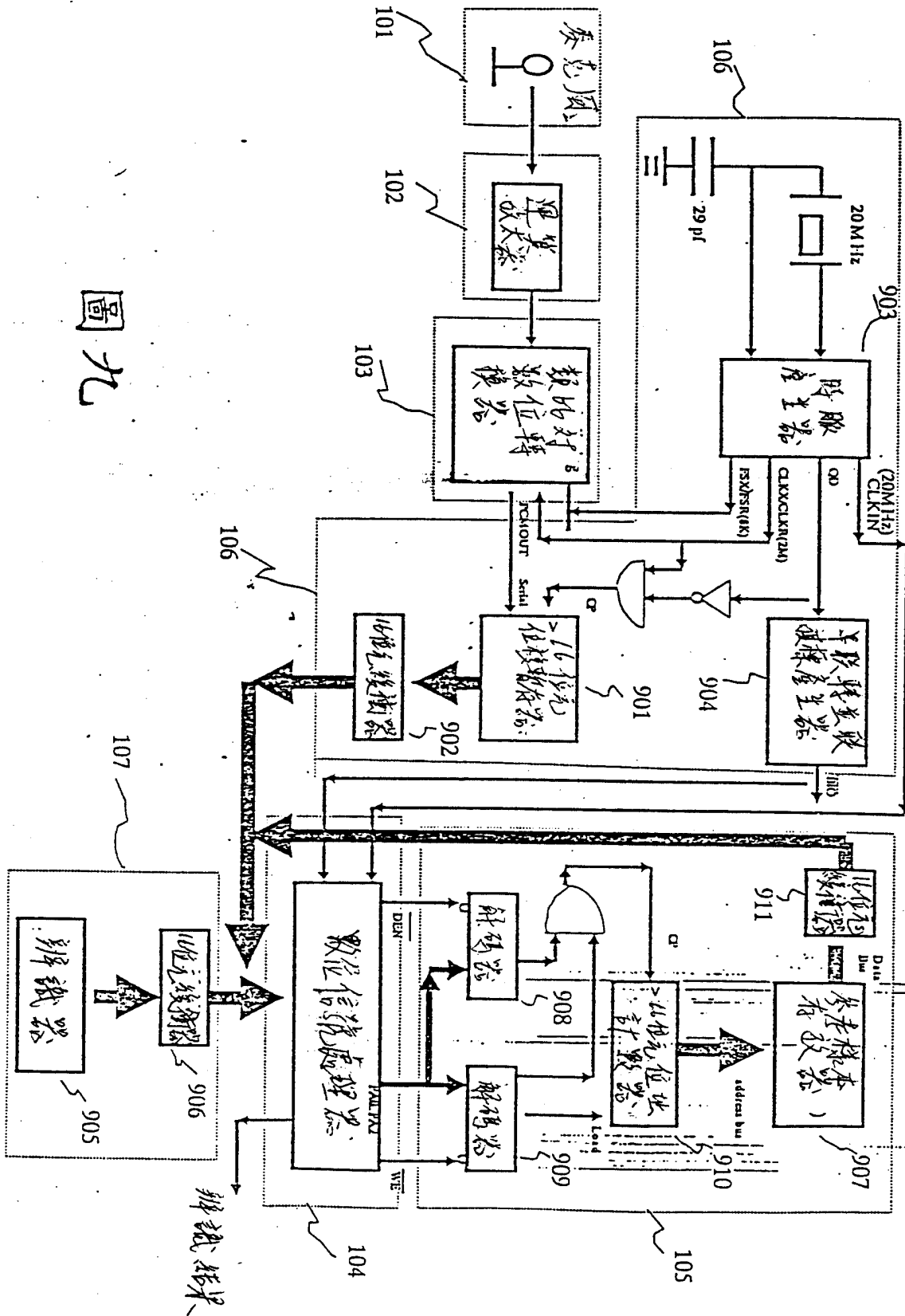
裝

訂

線

經濟部中央標準局員工消費合作社印製

圖式



圖九

(請先閱讀背面之注意事項再行繪製)

裝

訂

號