

發明專利說明書

(本說明書格式、順序及粗體字，請勿任意更動，※記號部分請勿填寫)

※申請案號：94139192

※申請日期：94.11.8

※IPC 分類：G06T 7/00(2006.01)

一、發明名稱：(中文/英文)

人類身體特徵姿勢資訊自動抽取系統/

Method for Automatic Key Posture Information Abstraction

二、申請人：(共 1 人)

姓名或名稱：(中文/英文)

國立交通大學/ National Chiao Tung University

代表人：(中文/英文) 張俊彥/ CHENG, CHUN-YEN

住居所或營業所地址：(中文/英文)

新竹市大學路 1001 號/

1001 Ta Hsueh Road, Hsinchu, Taiwan 300, R.O.C.

國 籍：(中文/英文) 中華民國/TW

三、發明人：(共 2 人)

姓 名：(中文/英文)

1. 廖弘源/ LIAO, HONG-YUAN

2. 陳敦裕/ CHEN, DUAN-YU

國 籍：(中文/英文)

中華民國/TW

四、聲明事項：

主張專利法第二十二條第二項 第一款或 第二款規定之事實，其事實發生日期為： 年 月 日。

申請前已向下列國家（地區）申請專利：

【格式請依：受理國家（地區）、申請日、申請案號 順序註記】

有主張專利法第二十七條第一項國際優先權：

無主張專利法第二十七條第一項國際優先權：

主張專利法第二十九條第一項國內優先權：

【格式請依：申請日、申請案號 順序註記】

主張專利法第三十條生物材料：

須寄存生物材料者：

國內生物材料 【格式請依：寄存機構、日期、號碼 順序註記】

國外生物材料 【格式請依：寄存國家、機構、日期、號碼 順序註記】

不須寄存生物材料者：

所屬技術領域中具有通常知識者易於獲得時，不須寄存。

五、中文發明摘要：

本發明之人類身體特徵姿勢抽取方法，係包括：由連續數位化畫面中抽取畫面中所含物件（object）之空間域特徵；利用機率計算方法，抽取該物件之形狀特徵；利用亂度計算方法偵測該連續畫面中所含之特徵姿勢資訊；除去冗餘之特徵姿勢；將所得之連串特徵姿勢與一碼書中所存之人類身體特徵姿勢樣型比對；及對配對之連串特徵姿勢進行編碼。

六、英文發明摘要：

The method for automatic key posture information abstraction of this invention comprises the steps of: Abstracting from a series of continuous digitized images spatial features of objects contained in said images; abstracting shape features of said objects using a method of probability calculation; detecting key posture information contained in said series of continuous images using a method of entropy calculation; removing redundant key postures; mating obtained key postures with key posture templates stored in a codebook; and encoding mated key postures.

七、指定代表圖：

(一)本案指定代表圖為：第 (1) 圖。

(二)本代表圖之元件符號簡單說明：

八、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

九、發明說明：

【發明所屬之技術領域】

本發明是關於一種電腦影像處理系統，特別是關於一種對數位化影片之畫面進行分析，自動在其中抽取人類身體特徵姿勢資訊之處理系統。

【先前技術】

對數位化影像檔中所含人類姿勢的分析乃是能夠成功分析數位化影像中所含之人類行動最重要的關鍵。分析數位影像中人類姿勢的困難主要有二。其一是人類身體的移動是一種分節（articulated）動作，因此對數位影像中「人類姿勢」的定義難度較高且較為複雜。其次，人類行為的特徵抽取，其實與對一系列影像畫面之處理相同。抽取其中包含的特徵時，資訊包括空間域資訊及時間域資訊。因而，以電腦系統進行數位影像之人類姿勢分析最大的挑戰，乃是如何在大量的時間域／空間域資料中正確抽取其特徵，以供其後進行比較與呼出。

過去，對於數位化影像中人類姿勢的分析技術，主要分成兩類，一為二維法，一為三維法。在二維方法中，Haritaoglu 等人曾提出一種 W4(what, where, when and who) 即時人體動作監控系統，透過對身體輪廓（silhouette）在垂直面及水平面之投影，以決定人類身體的各種姿態，例如：站、坐、彎、躺等。見 I. Haritaoglu 等人：「Ghost: A human body part labeling system using silhouettes」, Proc. Int. Conf. Pattern Recognitiore, 第 1 卷，第 77-82 頁，1998 年。

Bobick 與 Davis 則提出一種暫時性樣本建立方法，利用堆疊一系列之連續畫面而達成。該暫時性樣本可依據一連串畫面中之活動能量影像（motion energy images）及活動密集影像（motion intensity images），而抽取人類身體運動之特徵資訊。以運動為主的特徵資訊是由活動能量影像及活動密集影像中抽出，並供作

樣型比對之用。見 A. F. Bobick 與 J.W. Davis :「the recognition of human movement using temporal templates」, IEEE Trans. on Pattern Analysis and Machine Intelligence, 第 23 卷第 3 號, 2001 年 3 月。

在三維技術方面, Boulay 等人提出的方法乃是先計算移動像素在一參考軸上之投影, 而透過 PCA (Principle Component Analysis, 主成分分析) 學習二維之姿勢外形。其次則利用一三維之姿勢模型, 將該投影資訊抽離與攝影機所在位置的關係。見 B. Boulay 等人:「Human posture recognition in video sequence」, Proc. IEEE Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 第 23-39 頁, 2003 年。

Zhao 等人則提出一種三維人類身體模型, 用以檢驗所偵測到之影像移動區域是否為人類身體。其檢驗方法乃是透過一分節的人類行走模型來進行行走辨認。見 Zhao 等人:「Segmentation and Tracking of Multiple Humans in complex Situations」, Proc. IEEE. Int. Conf. On Computer Vision and Pattern Recognition, 第 2 卷, 第 8-12 頁, 2001 年。不過, 由於三維影像處理複雜, 且所需設備昂貴, 至今仍無適合於商業應用的三維人體活動分析技術。

為能有效建立對數位影像中人類行為自動且有效之分析系統, 首先必須能提供一種系統化及自動化的方法, 能在記錄人類行為的連續影像畫面中, 自動判別人類行為的特徵姿勢。但在已知之技術文獻中, 並未見有人對此技術提出過符合實用之發明, 而可系統化地自動在連續影像中抽取特徵姿勢資訊。

【發明內容】

本發明之目的乃在提供一種系統化的方法, 以自動在連續數位化影像中抽取人類身體之特徵姿勢資訊, 供作處理之用。

本發明之目的也在提供一種能在連續數位化影像中有效抽取人類身體特徵姿勢資訊的方法。

本發明之目的也在提供一種能簡化處理的人類身體特徵姿勢資訊抽取方法。

本發明之目的也在提供一種使用上述方法的人類身體特徵姿勢資訊抽取系統。

本發明之目的也在提供一種製作成本低廉的人類身體特徵姿勢資訊抽取系統。

依據本發明之人類身體特徵姿勢抽取方法，係包括：

由連續數位化畫面中抽取畫面中所含物件（object）之空間域特徵；利用機率計算方法，抽取該物件之形狀特徵；利用亂度計算方法偵測該連續畫面中所含之特徵姿勢資訊；除去冗餘之特徵姿勢；將所得之連串特徵姿勢與一碼書中所存之人類身體特徵姿勢樣型比對；及對配對之連串特徵姿勢進行編碼。

【實施方式】

雖不欲為任何理論所拘束，但一般認為，利用一連串之數位化影像畫面，可以記錄人類一段時間的身體活動。但由於所獲得之資料量過於鉅大，利用電腦系統針對所得之畫面，一頁一頁（frame）進行辨認及抽取特徵，並不可行。因此，透過建立「特徵姿勢」（key postures 或 feature postures）的方法，建立若干人類身體之特徵姿勢模型，並據以在影片之畫面中辨認其中所含之特徵姿勢，乃是最適用之作法。在建立特徵姿勢模型時必須滿足的條件包括：所提供的特徵姿勢種類不能太少，才能描述大部份人類身體的姿勢。其次，所提供的特徵姿勢種類也不能太多，否則將產生冗餘（redundancy）的問題，且其後之特徵處理步驟將極為費時。

為能適當建立及挑選畫面中的人類身體姿勢，本發明乃透過「亂度變動累積量」(cumulative entropy change) 之計算，作為選取特徵姿勢的指標，而在一記錄人類活動之影片中選取適當的人類身體特徵姿勢。選得之特徵姿勢再經由一亂度交叉比對步驟，以除去多餘的特徵姿勢。本發明透過上述方法，自一人體動作 (action, 指含有人類身體活動資訊的一段影片) 中選取合用的特徵姿勢。所選得之特徵姿勢再與一碼書 (code book) 中所含的特徵姿勢模型比對，進行特徵化或編碼。如此一來即可將一記錄人類活動的原始影片編碼成一串代表人類身體特徵姿勢的代碼。而在其後比對過程中，影像的比對乃能簡化成代碼的比對。提高處理效率及正確度。

第 1 圖表示本發明人類身體特徵姿勢資訊抽取方法一實施例流程圖。以下依據圖式說明本發明之實施例。如第 1 圖所示，本發明之人類身體特徵姿勢抽取方法主要包括數個步驟，亦即：由連續數位化畫面中抽取畫面中所含物件 (object) 之空間域特徵；利用機率計算方法，抽取該物件之形狀特徵；利用亂度計算方法偵測該連續畫面中所含之特徵姿勢資訊；除去冗餘之特徵姿勢；將所得之連串特徵姿勢與一碼書中所存之人類身體特徵姿勢樣型比對；及對配對之連串特徵姿勢進行編碼。

以下說明其詳細之步驟。唯需說明者，以下之詳細說明僅在例示本發明之實例，而不得用以限制本發明之方法。

參考第 1 圖。在利用本發明人類身體特徵姿勢資訊抽取方法抽取連續影像中的人類身體姿勢資訊時，首先於 (101) 取得一數位化影像檔，該數位化影像檔含有多數連續之數位化畫面。於 (102) 將每一畫面分割成多數方塊，於 (103)

對每一方塊計算其中所含之「活動像素」(active pixels, 指人體在每一方塊中所佔的像素)之密度。

在應用上, 可以不針對畫面之全部區域進行處理。例如, 透過習知之影像分析技術(例如背景消除 background subtraction 等方法), 可以在一畫面中找到該畫面所包含的物件。計算該物件之長寬比(aspect ratio), 找出包圍該物件影像之方塊。假設包圍該物件影像之方塊包含 $h \times w$ 個方格, h 代表高度方向之方格數, w 為寬度方向之方格數, 則影像方塊之活動像素密度 $P_i = N_j / N$; N_j 表在包圍該物件之方塊中, 第 i 方格之活動像素數量; N 表該方格內總像素數。

接著, 於(104)計算該連續畫面中影像物件之指數亂度。該影像物件之指數亂度(exponential entropy) H 可定義為:

$$H_{(t)} = \sum_{i=1}^{h \times w} p_i \exp(1 - p_i)$$

為提高處理效率, 以適合於即時處理之要求, 必須提供一種過濾機制, 以過濾不必要的影像畫面, 而取得最需要的影像。雖不欲為任何理論所限制, 但本發明人發現, 如一包圍影像物件之方塊中, 含有連結元素(connected components, 指相互連結在一起的活動像素)的方格數量較多, 則其該方塊之亂度將較高。事實上, 含有連結元素之方格數量, 與含有該等方格之方塊之亂度成正比。計算亂度的結果, 可以據以判斷該方塊之性質。

本發明係依據影像方塊之亂度(指數亂度)抽取影像物件之特徵, 亦即, 以其指數亂度代表影像物件之形狀資訊。在本發明中, 並將相連兩影像之人體姿勢之間距, 定義為兩影像物件之亂度(即兩影像方塊之亂度)之差值:

$$D_{\text{posture}}(t, t-1) = |H(t) - H(t-1)|$$

其中， $t, t-1$ 表影像（畫面）或身體姿勢在連續影像中之序數。

沿續以上定義，將特徵姿勢（key posture）定義為：當任一影像片段之累積亂度與前一影像片段之累積亂度差值大於一臨界值時，即視該影像片段所含的影像方塊為一特徵姿勢。該累積指數亂度差值 $C(P_p, P_Q)$ 可以下式算出：

$$C(P_p, P_Q) = \sum_{t=P}^{Q-1} D_{posture}(t, t-1);$$

P, Q 分別表自第 P, Q 畫框開始之一段連續影像， $P < Q$ 。

因此，於 (105) 依據上式計算該影片中各片段間之累積指數亂度差值 $C(P_p, P_Q)$ 。於 (106) 判斷各段代表特徵姿勢之連續影像片段。於 (107) 將各代表特徵姿勢之影像片段中任一影像取出，代表一特徵姿勢。在本發明之實例中，係選取第一張影像作為其特徵姿勢。當然，選用其他畫面或其他畫面之結合或處理後之結果，亦屬可行。

在處理時，也可先選定一畫面作為候選特徵姿勢。接著由該畫面開始統計影片（連續影像）之累積指數亂度差值。如累積亂度差值超過臨界值，則將該時點之畫面所示之姿勢，作為第二特徵姿勢。餘此類推。直到全部影片均處理完畢。此種方式可以便利即時處理。第 2 圖即顯示在 200 張連續影像畫面中，依據上述之累積指數亂度差值法所挑選出來的 18 張候選特徵姿勢。第 3 圖則為第 2 圖之來源影片各畫面之亂度（指數亂度）值示意圖。圖中，橫軸為畫面參數，縱軸為亂度值。

依上述方法所取得之影像，為可能的候選特徵姿勢畫面。但其中包括了多數冗餘的姿勢畫面，而可能在後級處理中浪費處理時間及資源。為消除冗餘的畫面，本發明也提供一種計算交叉指數亂度值計算法，計算不近似度 $d(P_p, P_Q)$ ，

用以去除不必要的畫面。

$$d(P_P, P_Q) = \sum_{i=1}^{h \times w} P_P(i) \exp(P_P(i) - P_Q(i)) + \sum_{i=1}^{h \times w} P_Q(i) \exp(P_Q(i) - P_P(i))$$

P_P, P_Q 為畫面。 $P_P(i), P_Q(i)$ 是第 i 個影像方塊之活動像素密度。

在應用上， P 與 Q 通常為僅含人體姿勢的畫面，因為背景已經去除，只剩下前景即人體的部分。計算時，對所有之畫面間之不近似度均計算乙次。如果 P 的方格數量與 Q 的數量不同，可以 P_P, P_Q 及其長寬比為依據各算一次。

於 (108) 計算出代表不同人體姿勢之畫面的不近似度 $d(P_P, P_Q)$ 。經過上述計算後，於 (109) 選出不近似度 $d(P_P, P_Q)$ 大於一臨界值之影像片段，作為冗餘的特徵姿勢影像，於 (110) 將該 P 影片片段所代表的候選特徵姿勢加以刪除。

第 4 圖即顯示將第 2 圖之候選特徵姿勢中之冗餘畫面刪除後之結果。圖中的黑框起之畫面，為不予刪除之畫面，代表選出之特徵姿勢。第 5 圖則顯示相對應之畫面之指數亂度值示意圖。

經過上述處理後之畫面，即為代表特徵姿勢之畫面，再於 (111) 與一特徵姿勢資料庫中所存之特徵姿勢樣型進行比對。比對時係計算特定特徵姿勢與資料庫中之特徵姿勢樣型之不近似度 D ，如下：

$$D(Q, T) = \sum_{i=1}^{h \times w} Q(i) \exp(Q(i) - T(i)) + \sum_{i=1}^{h \times w} T(i) \exp(T(i) - Q(i))$$

其中， Q 表一特徵姿勢畫面， T 表一樣型畫面， $D(Q, T)$ 表兩者之不近似度， $Q(i)$ 及 $T(i)$ 分別表畫面 Q, T 內所含之方格內之活動像素數量。

算出該特徵姿勢畫面與各樣型畫面之不近似度後，於 (112) 挑選不近似度最低之樣型畫面，作為其配對畫面。於 (113) 將配對成功之樣型之代碼給予該

特徵姿勢，完成對該特徵姿勢之編碼。

依上述步驟，對所有選定之特徵姿勢畫面均給予編碼，得到一串代碼。

將該一串代碼與一代表一種人類活動之連串動作之連串代碼進行比對。計算其近似度，即可判斷該影片片段所含資訊是否為含有人類活動之影像資訊，以及所代表之人類身體活動內容（種類）為何。

實施例

自一含有 6100 畫面（33 個片段）之影片中抽取人類身體特徵姿勢畫面。利用上述指數亂度計算及定義亂度計算方法，選出 44 張特徵姿勢畫面。第 6 圖即顯示所選出之 44 張特徵姿勢畫面。圖中，左上方之圖像為一行走中人體某一時點之姿勢，並顯示其相對應之特徵姿勢樣本為右方樣本集中第 7 號畫面（以粗框標示）。第 6 圖上方顯示該影片中人物在行走中，其特徵姿勢所對應之特徵姿勢樣本之編號及順序。

第 7 圖（a）顯示人類身體在行走過程中所呈現之所有特徵姿勢之順序。為特徵姿勢庫中所存之樣型。第 7 圖（b）顯示以上述方式所配對成功之特徵姿勢。圖中可見，前 3 張特徵姿勢畫面與樣本第 5 號比對相符；第 4、5 畫面則與第 6 號樣本比對相符。第 7 圖（c）顯示，特徵姿勢畫面中，第 5、6、7、8、38、39 共 6 張比對相符。透過本發明之編碼方式，可成功判斷該影片中之人物乃是在行走。

第 8 圖顯示本發明另一實施例之編碼過程。該實施例係對一人體由躺到起立之動作記錄影片，抽取特徵姿勢。其中，（a）顯示一個完整的由躺至起立之動作。（b）表示已比對符合之特徵姿勢畫面。（c）則顯示經過編碼之特徵姿勢。由本

實施例也可得知，本發明之方法可利用電腦迅速正確判斷一段影片中所含之人體活動種類。

以上是對本發明人類身體特徵姿勢自動抽取系統之說明。習於斯藝之人士，應可由以上說明得知本發明之精神，並據以作出各種變化與衍生。唯只要不脫離本發明之精神，均應包含於其申請發明專利範圍之內。

【圖式簡單說明】

第 1 圖表示本發明人類身體特徵姿勢資訊抽取方法一實施例流程圖。第 2 圖即顯示在 200 張連續影像畫面中，依據上述之累積指數亂度差值法所挑選出來的 18 張候選特徵姿勢。

第 3 圖則為第 2 圖之來源影片各畫面之亂度（指數亂度）值示意圖。第 4 圖即顯示將第 2 圖之候選特徵姿勢中之冗餘畫面刪除後之結果。

第 5 圖則顯示相對應之畫面之指數亂度值示意圖。

第 6 圖即顯示所選出之 44 張特徵姿勢畫面。

第 7 圖（a）顯示人類身體在行走過程中所呈現之所有特徵姿勢之順序。

第 7 圖（b）顯示以上述方式所比對成功之特徵姿勢。

第 7 圖（c）顯示特徵姿勢畫面中，第 5、6、7、8、38、39 共 6 張比對相符。

第 8 圖顯示本發明另一實施例之編碼過程。其中，（a）顯示一個完整的由躺至起立之動作。（b）表示已比對符合之特徵姿勢畫面。（c）則顯示經過編碼之特徵姿勢。

【主要元件符號說明】

十、申請專利範圍：

1. 一種人類身體特徵姿勢抽取方法，係利用一電子計算機從一連續數位化影像畫面中，選取含有代表人類身體特徵姿勢之資訊之畫面，分析該畫面所含代表人類身體特徵姿勢之資訊之方法，包括下列步驟：

由多數連續數位化畫面中，抽取畫面中所含物件（object）之空間域特徵；

利用機率計算方法，抽取該物件之形狀特徵；

利用亂度計算方法偵測該連續畫面中所含之特徵姿勢資訊；

除去冗餘之特徵姿勢；

將所得之連串特徵姿勢與一碼書中所存之人類身體特徵姿勢樣型比對；及

對配對之連串特徵姿勢進行編碼。

2. 如申請專利範圍第 1 項之人類身體特徵姿勢抽取方法，其中該抽取物件空間域特徵之步驟包括：

取得一組多數之連續數位化影像畫面；

將每一畫面分割成多數方塊；及

對每一方塊計算其中所含之「活動像素」之密度，該活動像素包括在每方塊中該物件所佔的像素。

3. 如申請專利範圍第 2 項之人類身體特徵姿勢抽取方法，其中該抽取影像物件特徵之步驟包括以影像方塊之亂度值代表影像物件之特徵。

4. 如申請專利範圍第 3 項之人類身體特徵姿勢抽取方法，其中之亂度為指數亂度。

5. 如申請專利範圍第 4 項之人類身體特徵姿勢抽取方法，其中該連續畫面中影像物件之指數亂度係以下式計算得出：

$$H_{(t)} = \sum_{i=1}^{h \times w} p_i \exp(1 - p_i)$$

其中， P_i 為活動像素密度， $P_i = N_i/N$ ； N_i 表在包圍該物件之方塊中，第 i 方格之活動像素數量； N 表該方格內總像素數。

6. 如申請專利範圍第 1 項之人類身體特徵姿勢抽取方法，其中當任一影像片段之累積亂度與前一影像片段之累積亂度差值大於一臨界值時，即視該影像片段所含的影像方塊為一特徵姿勢；該累積指數亂度差值 $C(P_p, P_Q)$ 以下式算出：

$$C(P_p, P_Q) = \sum_{t=p}^{Q-1} D_{posture(t, t-1)};$$

其中， P_p 指一長度為時段 P 之連續畫面中所含的人類身體姿勢。

7. 如申請專利範圍第 1 項之人類身體特徵姿勢抽取方法，其中選定特徵姿勢資訊之步驟包括：

先選定一畫面作為候選特徵姿勢；

接著由該畫面開始統計該連續畫面之累積指數亂度差值；

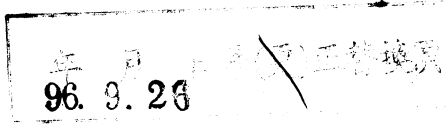
如累積亂度差值超過臨界值，則將該時點之畫面所示之姿勢，作為第二特徵姿勢；及

餘此類推，直到全部畫面均處理完畢。

8. 如申請專利範圍第 1 項之人類身體特徵姿勢抽取方法，其中除去冗餘特徵姿勢之步驟包括：

計算相鄰畫面片段的不近似度 $d(P_p, P_Q)$ ；

選出與前段影像片段不近似度大於一臨界值之影像片段，作為冗餘的特徵姿勢影像；及



將該 P 影片片段所代表的候選特徵姿勢加以刪除。

9. 如申請專利範圍第 8 之人類身體特徵姿勢抽取方法，其中之不近似度 $d(P_P, P_Q)$ ，係以下適計算而得：

$$d(P_P, P_Q) = \sum_{i=1}^{n \times w} P_P(i) \exp(P_P(i) - P_Q(i)) + \sum_{i=1}^{n \times w} P_Q(i) \exp(P_Q(i) - P_P(i))$$

$P_P(i)$ 是第 i 個影像方塊， P 和 Q 為畫面。

10. 如申請專利範圍第 1 項之人類身體特徵姿勢抽取方法，其中之樣型比對步驟包括：

將代表特徵姿勢之畫面與一特徵姿勢資料庫中所存之特徵姿勢樣型進行比

對；

計算其不近似度 D ；及

挑選不近似度最低之樣型畫面，作為其配對畫面。

11. 如申請專利範圍第 10 項之人類身體特徵姿勢抽取方法，其中之不近似度 D 係以下式計算而得：

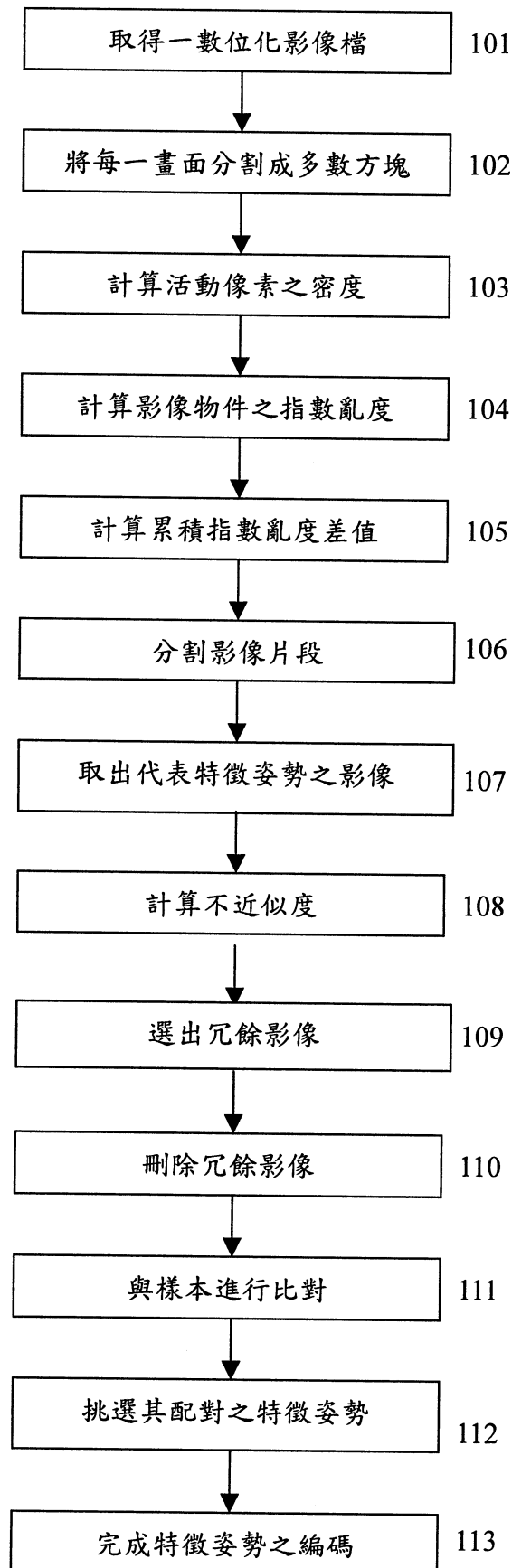
$$D(Q, T) = \sum_{i=1}^{h \times w} Q(i) \exp(Q(i) - T(i)) + \sum_{i=1}^{h \times w} T(i) \exp(T(i) - Q(i))$$

其中， Q 表一特徵姿勢畫面， T 表一樣型畫面， $D(Q, T)$ 表兩者之不近似度，

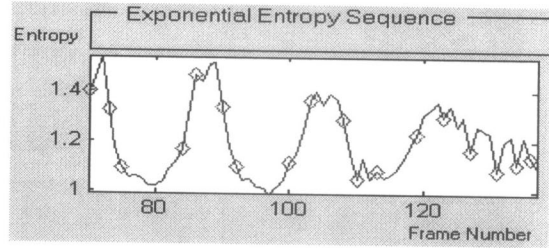
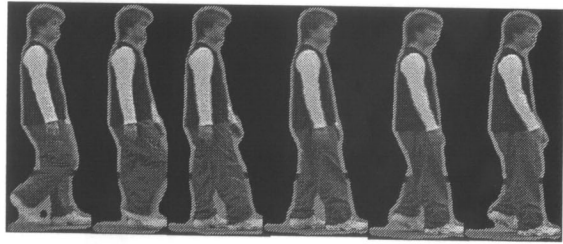
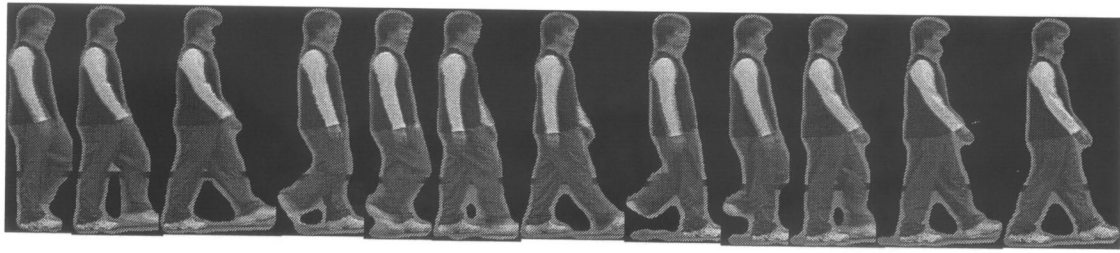
$Q(i)$ 及 $T(i)$ 分別表畫面 Q ， T 內所含之方格內之活動像素數量。

12. 如申請專利範圍第 10 項之人類身體特徵姿勢抽取方法，另包括一將所得之該連續畫面之代碼與一代表一種人類活動之連串代碼進行比對，計算其近似度及判斷該影片片段所含資訊是否為含有人類活動之影像資訊，及其種類之步驟。

十一、圖式：

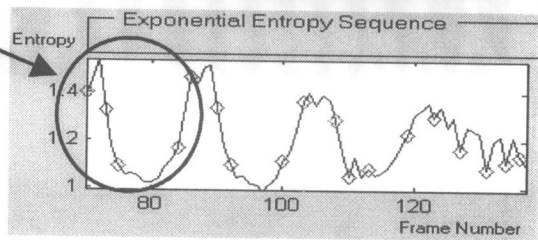
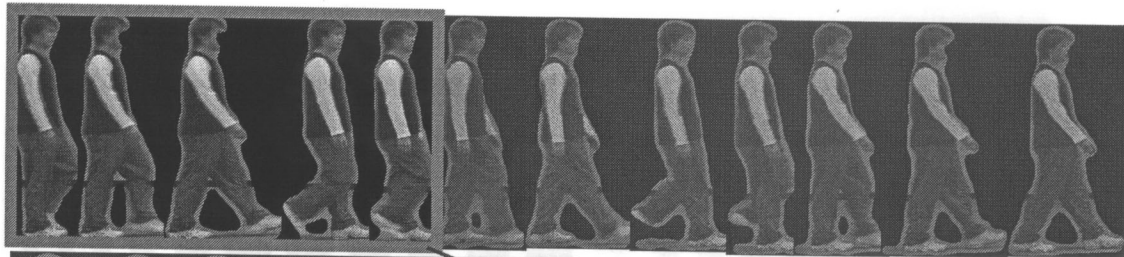


第 1 圖



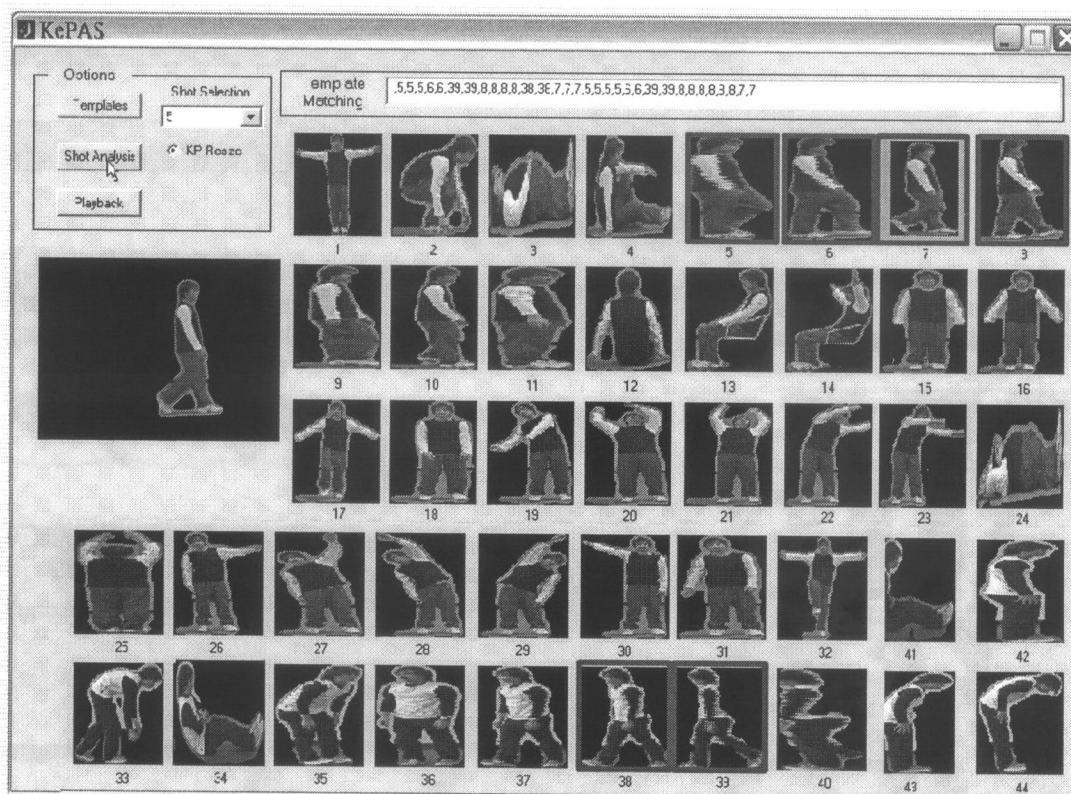
第 2 圖

第 3 圖

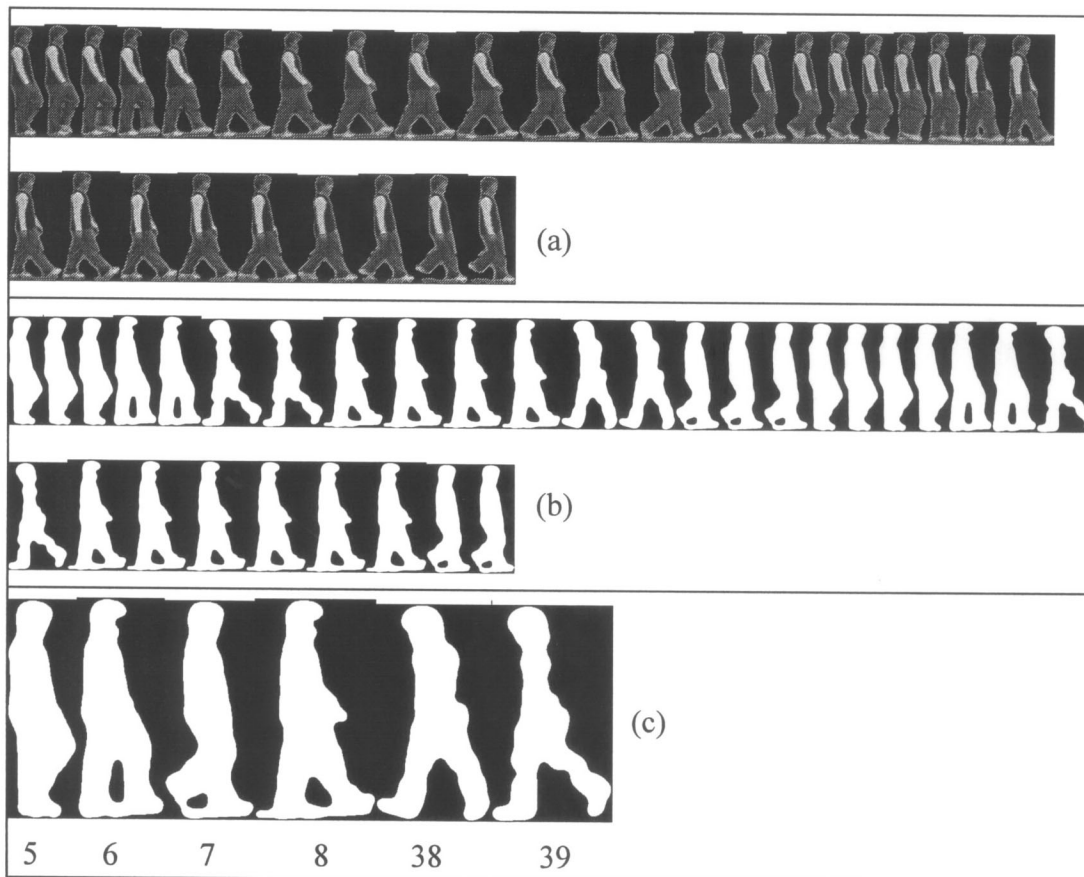


第 4 圖

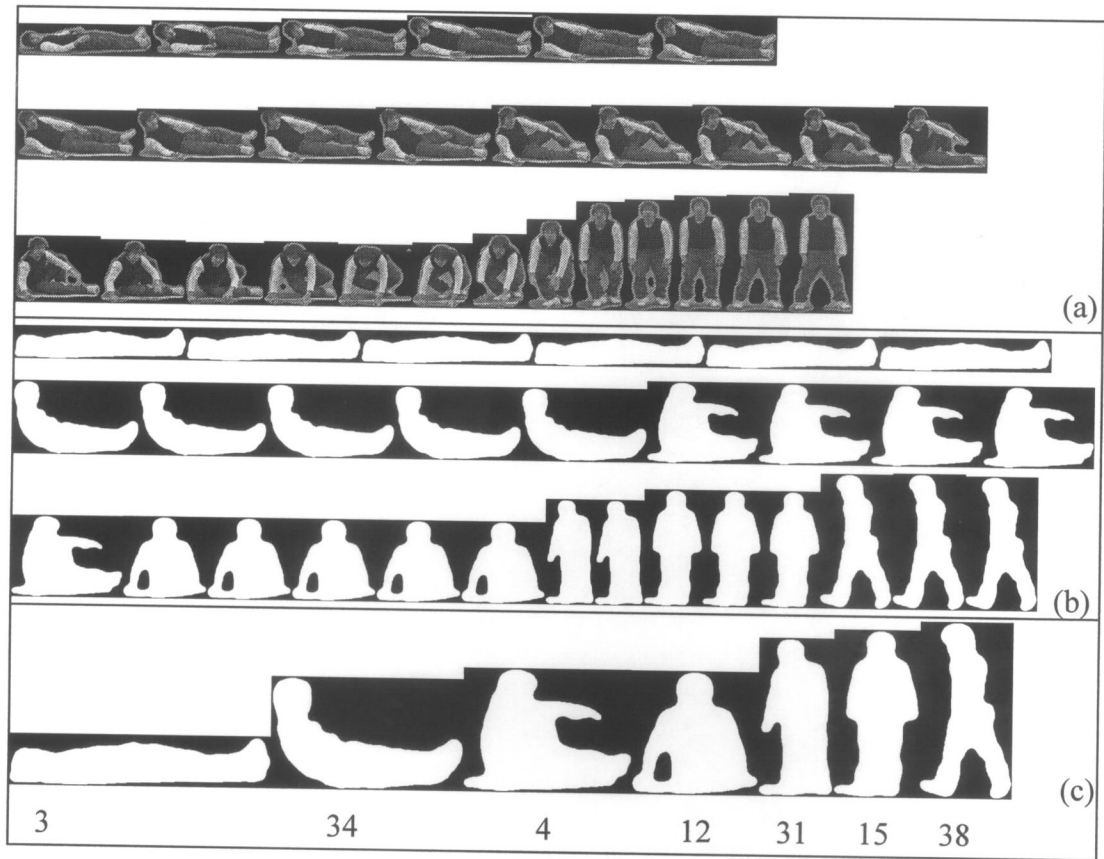
第 5 圖



第 6 圖



第 7 圖



第 8 圖