彩色人體影像之架構抽取

Framework Extraction for a Color Human Body Image[1]

陳永盛[2]　　邱淑姝
Yung-Sheng Chen[2] and Shu-Shu Chiu

元智大學電機工程學系
Department of Electrical Engineering
Yuan Ze University
Taoyuan, Taiwan, Republic of China

摘　要

從影像中抽取人體架構是電腦視覺領域中的一個重要課題。在一個穿著長袖與長褲的前視站立人像條件下，我們提出一系統方法可以完成抽取工作。為了能有效地抽取人體架構，根據解剖學的觀點，人體相關部位與其間的關係首先被定義出來。我們方法的三個主要步驟是：人臉偵測、衣服與褲子區域的分割、以及主要人體部位的定位。人臉偵測是藉著人臉膚色抽取與人臉模版比對來完成。一個色彩紋理相似度量測被設計來分割衣服與褲子區域。最後，我們從相似度量測中導出同質與異質資訊，將其結合人體各部位的關係後依序定出人體之軀幹、手臂、臀部、以及腿部區域。實驗已證實我們方法的可行性以及可應用性。

關鍵詞：人臉偵測、人體架構、色彩紋理分割、同質資訊、異質資訊

Abstract

Extracting human body framework from images is a significant topic in computer vision field. Based on an assumption: only one front-viewed standing human body dressing clothes with long sleeves and long trousers, a framework extraction approach is presented. The main parts and their relationships considered in a human body are defined first from the viewpoint of anatomy. Three main steps devised in our approach are face detection, segmentation of clothes and trousers regions, and positioning main human body parts. The face detection is performed using face skin color extraction and face template matching. A color texture similarity measure is designed for the segmentation of clothes and trousers. The homogeneous and inhomogeneous information are finally derived from similarity measurements and combined with the defined relationships of a human body framework to locate the trunk region, arm parts, hip region, and leg parts in order. Experiments have confirmed the feasibility and applicability of the proposed approach.

Keyword: face detection, human body framework, color texture segmentation, homogeneous information, inhomogeneous information.

---

[1] This work was supported by National Science Council under Grant NSC89-2218-E-155-001.

[2] To whom all correspondence should be addressed.

1. Introduction

In order to enhance the human-computer interaction (HCI), many researches of computer vision have been conducted in the field of gesture recognition [1, 2], face recognition [3, 4], and expression recognition [5, 6]. Recently, some results from these have been moved into commercial products one after another, e.g., the Magic Board [7] and the person identification [8]. In the meantime, it is not unexpected that these researches have intuitively and immediately turned to the hot topic of researching human body. The topic of "looking at people," which gives computer the ability to detect, track, identify people, and further interpret human behavior, has played an important role in computer vision field [9, 10]. In this study, the related researching literatures may be classified into four classes based on the type of sources, i.e., static 2D image [11-13], static 3D image [14], dynamic 2D image frames [15-21], and dynamic 3D image frames [22-24]. Except for these literatures, a good survey for the visual analysis of human movement [25] may also help us to understand the history of this field.

These literatures reveal that most of researches are focusing on the study of human movements because the human posture is easier to be detected and analyzed due to the variation between consequent frames. However, if only a static single image (This is a real case since the photograph is often used in our daily life) is adopted, the posture extraction will become difficult due to the lack of temporal information compared with a sequence of image frames. In addition, it is worthy of exploring the image processing techniques used to analyze the static human body image. The goal of this paper is therefore to present our recent research on the extraction of human body framework from a color image. Some interesting and valuable problems will also be pointed out.

The name of "human body framework," which may be abbreviated as HB-framework in this paper, is sometimes called "human body model" or "human stick figure." The more precise definition of the human body framework may be called as "human body skeleton" from the viewpoint of anatomy. However, because the head (face), trunk, and hip are the main parts of the human body, the "human body framework" will be used in our current research.

The concept of our HB-framework extraction comes from that of thinning (or skeletonization) [26, 27], which is a process of successively stripping the boundary of a figure until a connected, single-pixel framework or skeleton remains. The found lines and junctions from a thinned image as well as the relationships among them [28] are usually applied to pattern recognition such as optical character recognition (OCR). Accordingly, it is our main research motivation to develop a systematic approach for HB-framework extraction, which may be the main and useful primitives for further human body recognition and analysis.

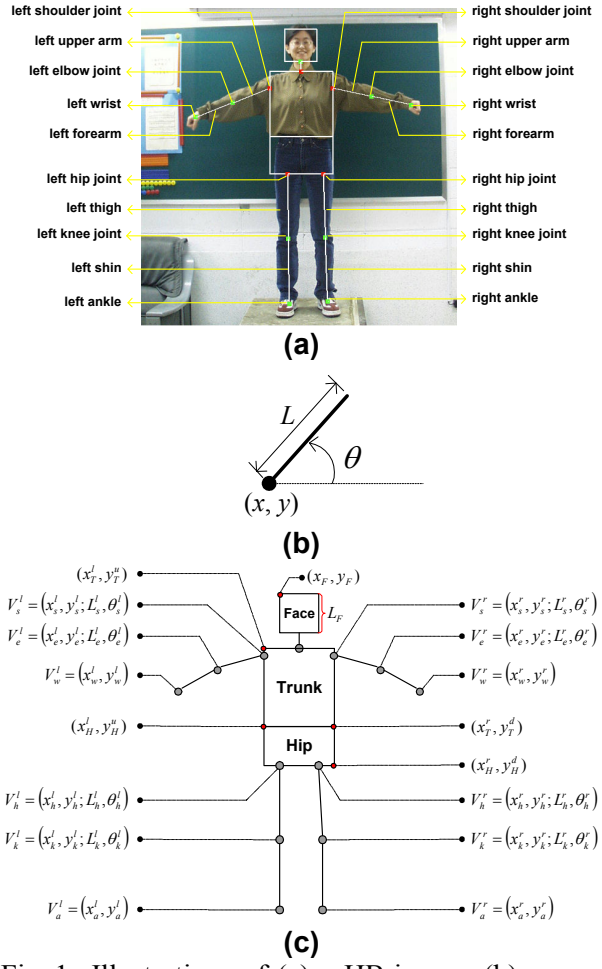


Fig. 1. Illustrations of (a) a HB-image, (b) a vector format $V=(x,y;L,\theta)$, and (c) data formats defined in our framework.

2. A human body framework

Before presenting our algorithms to extract the framework from a color human body (HB) image, the names of the body parts are displayed in Fig. 1(a) from the viewpoint of anatomy. The main constraint of our current work is confined to

deal with a front-viewed static HB-image having clothes and trousers regions. The concerned parts of a HB-image include face; trunk; hip; and left/right parts of shoulder joint, upper arm, elbow joint, forearm, wrist, hip joint, thigh, knee joint, shin, and ankle. For the sake of displaying the final extracted HB-framework, it is suitable to adopt a rectangle for representing the face, trunk, and hip region; and adopt a vector format (as illustrated in Fig. 1(b)) for representing the other body parts, as indicated in Fig. 1(c).

From the real image shown in Fig. 1(a), three main steps are developed in this paper to extract the HB-framework. They are (1) the face detection (see Section 3), (2) the segmentation of clothes and trousers regions (see Section 4), and (3) the positioning of all human body parts for accomplishing HB-framework extraction (see Section 5) according to the relationships of basic framework defined in Fig. 1(c).

3. Face detection

Several face detection algorithms have been proposed. Two major characteristics, face skin color and face template, are adopted in our face detection due to the use of color image.

3.1. Analysis of face skin color

To analyze the distribution of face skin color, the hue component in a color image is used and the concept of skin color distribution model proposed by Wu et al. [29] is adopted. The hue distribution of face skin color exists a hue range ($[hue^l_{skin}, hue^r_{skin}]$) of face skin color such that the probability $p(h) < 1$, where $h \in [hue^l_{skin}, hue^r_{skin}]$. In our experiments, a set of face skin regions is selected from thirty color images. The skin color is then converted from RGB color space into LCH color space. The probability density function of hue for all pixels in the set of selected face regions is finally constructed to determine the hue range. In this study, $[hue^l_{skin}, hue^r_{skin}] = [6,\ 61]$.

3.2. Construction of gray face template

Because a region having hue value $h \in [hue^l_{skin}, hue^r_{skin}]$ depends on the given color image, it is necessary to use an advanced procedure to detect the face location. After extracting the hue information of face skin color, the color pixels addressed by hue information will be converted into the gray ones (called face-skin-color pixels) and others denoted by black ones (called non-face-skin-color pixels). Thus the final face region may be detected directly by convolving a defined gray face template onto the region having face-skin-color pixels. The gray face template ($\mathbf{T}_{\text{gray face}}$) may be constructed as follows.

(1) Take thirty gray images of the human face from the database [30], average the gray information for all images at each pixel location, and obtain a mean gray face image.
(2) Smooth the mean gray face image by a Gaussian function, and manually locate the face portion using a curve fitting method [31]. The gray face template $\mathbf{T}_{\text{gray face}}$ is then obtained.

(a) (b)

Fig. 2. (a) The gray image. (b) The final detected face region.

3.3. Face detection algorithm

After accomplishing the analysis of face skin color and the construction of gray face template, let the original color image shown in Fig. 1(a) be an input one for illustration, we present the face detection algorithm as follows.

(1) Convert the input image from RGB color space into the LCH color space.
(2) Use the known hue range $[hue^l_{skin}, hue^r_{skin}]$ = [6, 61] found in Section 3.1 to extract the possible pixels constituting the possible face regions from the LCH color image.
(3) Convert the color image into the gray image as shown in Fig. 2(a). Then convolve it by using the known gray face template $\mathbf{T}_{\text{gray face}}$ from smaller size to larger size. Note here that because the face region size is usually unknown, it is necessary for the convolution by using the different template size. In our experiments, the range of template size is adjusted from 40×40 to 64×64.

(4) After the convolution, the location having peak value and the used template size will be reported. In this example, the location is (232, 35), and the template size is 52×52. Based on the data format defined in HB-framework, we denote it by $(x_F, y_F; L_F)$ = (232, 35; 52) and the detection result is shown in Fig. 2(b).

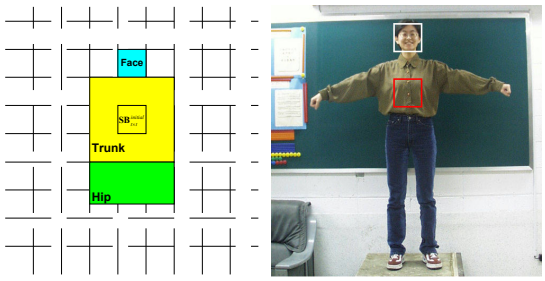


Fig. 3. (a) The size ratios among face, trunk, and hip used in our study, where $\mathbf{SB}_{t\times t}^{initial}$ denotes the initial seed block. (b) The initial seed block $\mathbf{SB}_{52\times 52}^{initial}$ assigned will be the first step to determine the final seed block and texture features.

4. Primary segmentation

Color texture in the clothes and trousers regions is an apparent cue, and is suitably used for the segmentation of these regions. To effectively segment a region having the same color texture from an image, a seed block containing enough texture information should be determined first [32], where the seed block is denoted as $\mathbf{SB}_{t\times t}$. The texture information includes color co-occurrence matrix (**CCM**) and color occurrence vector (**COV**), which will be regarded as the texture features and be defined later. Of course, the initial seed block, $\mathbf{SB}_{t\times t}^{initial}$, should be assigned first onto a suitable location and region. According to the size ratio among face, trunk, and hip of human body painting, we may define the trunk and hip regions as the function of face size (L_F), i.e., $L_{T(width)} = L_{T(height)} = 3L_F$ for the trunk region, $L_{H(width)} = 3L_F$ and $L_{H(height)} = 1.5L_F$ for the hip region, and shown in Fig. 3(a). Since the clothes region is segmented first, for the further determination of the final seed block and texture features, we assign the initial seed block $\mathbf{SB}_{52\times 52}^{initial}$ in the trunk region as shown in Fig. 3(b) by the face size ($L_F = 52$) which has been obtained in the face detection stage. After evaluating the texture similarity by the texture features for all indexed blocks with the size $t \times t$ in the image, the region **S** possessing the texture features (**CCM** and **COV**) will be highlighted and segmented.

4.1. Hue quantiztion

The hue component ranged from 0° to 360° is used to construct a histogram for quantization. A so-called circular histogram is used to join 0° and 360° together for a hue histogram [33]. Because a histogram often contains many rugged peaks, this will result in a difficulty with thresholding. Hence, a scale-space filter [34] is used to smooth the histogram. After scale-space filtering, we can easily find valleys in the histogram by computing the first and second derivatives of the histogram. Then the *k*-means algorithm is further applied to accomplish the classification task [35]. If the image is quantized to *k* hue values, the quantized colors will be indexed as {1, 2, …, *k*}. For example in Fig. 4, here the two valleys (thresholds) are 23 and 85, and three obtained clusters imply three quantized colors, i.e., $k = 3$.

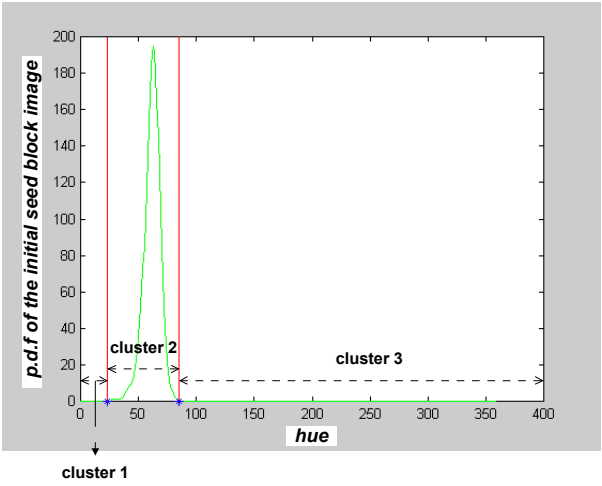


Fig. 4. The hue quantized result for the initial seed block image $\mathbf{SB}_{52\times 52}^{initial}$ shown in Fig. 3(b).

4.2. Texture features

Since only the hue component is used for the quantized colors, the gray level co-occurrence matrix may be adopted for the texture analysis [32, 36], and named as color co-occurrence matrix, **CCM**. Assume the block image is of size $t \times t$.

Let G be the set of k quantized color indexes and g represent the color index function of the image block. Let also the $(i,j)th$ entry of the **CCM** record the occurrence frequency that a pixel (m,n) with the ith color (denoted by $g(m,n)=i$) has an 8-connectivity neighborhood (p,q) with the jth color (denoted by $g(p,q)=j$), where (m,n), $(p,q)\in(t\times t)$, and $i,j\in G$. We define a test function: $\delta_{ij}(m,n;p,q)=1$ if (m, n) and (p, q) belong mutually to be an 8-connectivity neighborhood; otherwise $\delta_{ij}(m,n;p,q)=0$. Thus each entry of **CCM** normalized in the range of [0, 1] may be expressed by

$$CCM(i,j)=\left[\sum_{\forall(m,n),(p,q)\in(t\times t)}\delta_{ij}(m,n;p,q)\right]/(t\times t) \quad (1)$$

Another useful information is color occurrence vector **COV**. Each entry of **COV** records the frequency that ith color occurs (f_i) in the image and defined by

$$COV(i)=f_i/(t\times t) \quad (2)$$

Based on the result of quantized colors (k = 3 for example, refer to Fig. 4) for the block image $\mathbf{SB}^{initial}_{52\times52}$ shown in Fig. 3(b), we have G = {1, 2, 3}. Its **CCM** and **COV** can be obtained as

$$\mathbf{CCM}=\begin{bmatrix}0.00000 & 0.00000 & 0.00000\\ 0.00000 & 0.00000 & 0.00034\\ 0.00000 & 0.00034 & 0.99931\end{bmatrix} \text{ and}$$

$$\mathbf{COV}=\begin{bmatrix}0.00000 & 0.00000 & 0.999669\end{bmatrix}.$$

These two texture features with the initial block size of 52×52 will be further used to adaptively determine the final seed block $\mathbf{SB}^{final}_{t\times t}$, whose self-contained texture features will be finally used to segment the desired region.

4.3. Determination of final seed block

So far only the texture features **CCM** and **COV** of the initial seed block $\mathbf{SB}^{initial}_{t\times t}$ have been found, but it is unknown whether these features are stable or not. The stability of texture features considered here may be described as follows. Since we cannot exactly state that the texture features of $\mathbf{SB}^{initial}_{t\times t}$ contain enough information, it is reasonable to increase the block size until the texture features are stable. Next, to confirm the little variation of texture features involved in the selected block, it is also suitable to decrease the block size until the texture features are out-of-stable. The block size before instability of texture features occurs is just the wanted final seed block $\mathbf{SB}^{final}_{t\times t}$. To evaluate the similarity of texture features between two block images $\mathbf{SB}_1$ and $\mathbf{SB}_2$, two distances are defined respectively for **CCM** and **COV**,

$$D_{\text{ccm}}=\sum_{i=1}^{k}\sum_{j=1}^{k}\left|CCM_1(i,j)-CCM_2(i,j)\right|, \quad (3)$$

$$D_{\text{cov}}=\sum_{i=1}^{k}\left|COV_1(i)-COV_2(i)\right| \quad (4)$$

Note here that once k is found from the hue quantization of $\mathbf{SB}^{initial}_{t\times t}$ image, e.g., k = 3 with two thresholds in the current illustration. The given thresholds for k-clustering will be used to extract the texture features for any block image during the segmentation stage. Since the scales of D_{ccm} and D_{cov} are different, and to increase the fuzzy similarity, a Z-type standard function is defined by the following equation,

$$Z(x;a,b)=\begin{cases}1, & x\le a,\\ 1-\dfrac{2(x-a)^2}{(b-a)^2}, & a<x\le\dfrac{a+b}{2},\\ \dfrac{2(x-b)^2}{(b-a)^2}, & \dfrac{a+b}{2}<x\le b,\\ 0, & b<x.\end{cases} \quad (5)$$

We use the Z-type function to transform the distance into the similarity measurement. For the measurement, the smaller distance represents the higher similarity. Thus the similarity (S) of texture features between two block images may be evaluated by

$$S=e^{-[1-Z(D_{\text{ccm}};a,b)\times Z(D_{\text{cov}};a,b)]^2} \quad (6)$$

In our experiments, we choose (a, b) to be (0.05, 0.5). If the texture features between two blocks are very similar, $D_{\text{ccm}}\to0$ and $D_{\text{cov}}\to0$. This implies $Z(D_{\text{ccm}};a,b)\to1$ and $Z(D_{\text{cov}};a,b)\to1$, thus the similarity S will approach to 1, and *vice versa*.

Based on the stability consideration of texture features and the defined similarity function S, we construct the following procedure to determine the final seed block $\mathbf{SB}^{final}_{t\times t}$ having the stable texture features.

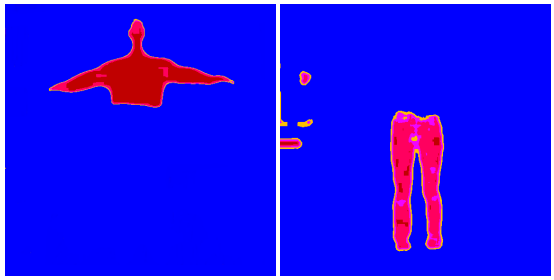
(1) Initialize the seed block $\mathbf{SB}_{t\times t} = \mathbf{SB}_{t\times t}^{initial}$, where the size length t is given, e.g., $t = L_F$.
(2) Take the eight nearest neighboring image blocks (denoted by $\mathbf{NB}_1 \sim \mathbf{NB}_8$) of $\mathbf{SB}_{t\times t}$ with the same size of $\mathbf{SB}_{t\times t}$, then compute (**CCM**)s and (**COV**)s for all the nine blocks.
(3) Compute the similarity (S_i) between $\mathbf{SB}_{t\times t}$ and $\mathbf{NB}_i$, $i = 1,2,...,8$. Thus there are eight similarity measurements.
(4) If the total number of $S_i > 0.8$ is smaller than 6, $t \leftarrow t+1$ and go to step 2. Otherwise go to step 5.
(5) Take the image block $\mathbf{SB}_{(t-1)\times(t-1)}$, compute its (**CCM**) and (**COV**), then compute the similarity S between $\mathbf{SB}_{t\times t}$ and $\mathbf{SB}_{(t-1)\times(t-1)}$.
(6) If $S > 0.8$, $t \leftarrow t-1$ and go to step 5. Otherwise stop this procedure.

The output of the above procedure will be the final seed block $\mathbf{SB}_{t\times t}^{final}$ containing the desired texture features $\mathbf{CCM}^{final}$ and $\mathbf{COV}^{final}$. Fig. 5 shows the found final seed block $\mathbf{SB}_{21\times 21}^{final}$ by applying this procedure to the initial seed block $\mathbf{SB}_{52\times 52}^{initial}$ given in Fig. 3(b).

Fig. 5. The found final seed block $\mathbf{SB}_{21\times 21}^{final}$.

4.4. Segmentation results

The segmentation scheme is performed as follows. First create a null image **S** having the same size as the original one to store the measuring results of texture similarity, and give the found seed block $\mathbf{SB}_{t\times t}^{final}$ possessing texture features $\mathbf{CCM}^{final}$ and $\mathbf{COV}^{final}$. For each location of the input image, take its $t\times t$ block image and compute the texture similarity between the block and $\mathbf{SB}_{t\times t}^{final}$, and record the computed similarity in the corresponding location of **S**.



(a) (b)

Fig. 6. Segmentation results of (a) clothes region ($\mathbf{S}_{clothes}$) and (b) trousers region ($\mathbf{S}_{trousers}$) for the HB-image shown in Fig. 1(a).

Along the current illustration for the clothes region, the segmented result $\mathbf{S}_{clothes}$ is shown in Fig. 6(a) after performing the above segmentation algorithm. Similarly, based on the size ratio illustrated in Fig. 3(a), a first seed block may also be roughly assigned on the hip region, then the trousers region $\mathbf{S}_{trousers}$ shown in Fig. 6(b) can be obtained by performing the presented segmentation algorithm. In accordance with the results, some other regions possessing the determined texture features but not belonging to a true clothes region or a true trousers region may also be highlighted. This will not influence the further HB-framework extraction since the necessary trunk region and hip region have been involved in the found $\mathbf{S}_{clothes}$ and $\mathbf{S}_{trousers}$, respectively. In next section, the information contained in $\mathbf{S}_{clothes}$ and $\mathbf{S}_{trousers}$ will be further processed to facilitate the extraction of the HB-framework.

5. HB-framework extraction

Fig. 6 shows two facts of the found similarity information of texture features: (1) the desired region is connected, and (2) there are many folds or shadows, in particular, appearing at the boundary of the region. These facts are useful to position the parts of a HB-framework by means of exploring the homogeneous and inhomogeneous properties in a local area.

5.1. Information development

In information theory, the definition of information is concerned with the probability of occurrence of the various messages (say source symbols). The entropy depends on the probability of the various source symbols. Let a zero-memory

information source with n symbols be defined by the set $S = s_j,\ j = 1,2,...,n,$ and the probabilities $P(s_j) = P_j,\ j = 1,2,...,n.$ The entropy of the set of source symbols may be expressed by

$$H(S) = \sum_{j=1}^{n} P_j \log P_j^{-1}. \tag{7}$$

The maximum value of the entropy is achieved if and only if all the source symbols are equiprobable.

We apply the principle of the maximum entropy to explore *HI* and *IHI* for any pixel in the image containing similarity information within a rectangular neighborhood ranged by (2*i*+1)×(2*i*+3). That is, we can find an effective rectangular range $RR(i_e)$, for a pixel (x,y), which contains the maximal information for the pixel as the two expressions

$$HI_{RR(i_e)}^{(\max)}(x,y) = \max_{\forall i} HI_{RR(i)}(x,y), \tag{8}$$

for homogeneous computation,

and

$$IHI_{RR(i_e)}^{(\max)}(x,y) = \max_{\forall i} IHI_{RR(i)}(x,y), \tag{9}$$

for inhomogeneous computation.

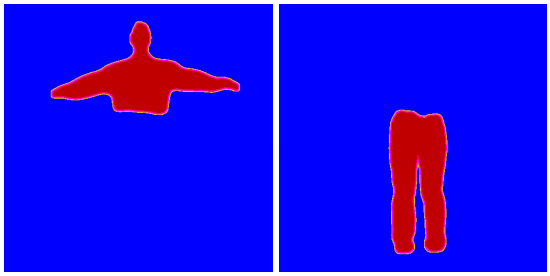
For computing *HI*, given a specified rectangular range $RR(i)$, let $\sigma_j = \sum_{\forall (m,n)\in(2j+1)\times(2j+3)} S(m,n)$ be the sum of similarity measurements for all pixels in $RR(j)$, and $P_{RR(0)}$ be the initial probability that σ_0 appears in $RR(0)$. The relative probability $P_{RR(j)}$ of having σ_j in $RR(j)$ is defined as

$$P_{RR(j)} = P_{RR(0)}/(2j+1)(2j+3),\quad \text{for } 0 < j \le i, \tag{10}$$

and $\sum_{j=0}^{i} \sigma_j P_{RR(j)} = 1$, where the normalizing factor $1/(2j+1)(2j+3)$ represents only the relative contribution of j to the defined probability and is not a unique selection. Based on these definitions, *HI* is modified from (7) and given by

$$HI_{RR(i)}(x,y) = \sum_{j=0}^{i} \sigma_j (P_{RR(j)} \log P_{RR(j)}^{-1}) \left(C(j+1)^{-1}\right), \tag{11}$$

where $C/(j+1)$ approximates to the normal distribution, and directly multiplies the corresponding self-information $\sigma_j(P_{RR(j)} \log P_{RR(j)}^{-1})$ to simulate the significance of the pixel decreased from the center to the outside in the considered range. The constant C = 1.5 is used in our experiments.



(a) (b)

Fig. 7. The clean results of (a) clothes region ($\mathbf{I}_{clothes}^{HI}$) and (b) trousers region ($\mathbf{I}_{trousers}^{HI}$).

Based on (8), for each pixel, there exists an effective index i_e whose $RR(i_e)$ contains the maximal homogeneous information $HI_{RR(i_e)}^{(\max)}$. We use an image $\mathbf{I}^{HI}$ having the same size as the original one to store the effective index i_e of the *HI*. The smaller i_e represents the stronger HI-intensity for the concerned similarity information. After performing (8), the corresponding $\mathbf{I}_{clothes}^{HI}$ and $\mathbf{I}_{trousers}^{HI}$ are obtained and illustrated in Fig. 7(a) and (b), respectively. As expected, the results are quite clean for further HB-framework extraction.

Because there are many folds or shadows, in particular, appearing at the boundary of $\mathbf{S}$, it is necessary to compute *IHI*. In our study, all the designed rules for *IHI* are the same as those for *HI*, except for the definition of $\sigma_j = \sum_{\forall (m,n)\in(2j+1)\times(2j+3)} \left((1-S(m,n))^{1/2} + S(m,n)/2\right)$. Similarly, let $\mathbf{S}$ containing the similarity information be an input image. For each pixel, there exists an effective index i_e whose $RR(i_e)$ contains the maximal inhomogeneous information $IHI_{RR(i_e)}^{(\max)}$. When the most of similarity values $S(m,n) \to 0$ or $S(m,n) \to 1$ in $RR(j)$, this will enable i_e to be large, whereas the similarity varies much in $RR(j)$ will make i_e small. That is, the smaller i_e will represent the stronger IHI-intensity for the concerned similarity information. Thus the boundary of $\mathbf{S}$ may be highlighted in $\mathbf{I}^{IHI}$. Fig. 8 shows the result of $\mathbf{I}_{clothes}^{IHI}$, which will be used to estimate the initial region of trunk. In our approach, we do not find $\mathbf{I}_{trousers}^{IHI}$ since the hip region and leg parts may be sequentially derived in accordance with the obtained trunk region and the defined relationships of a HB-framework.

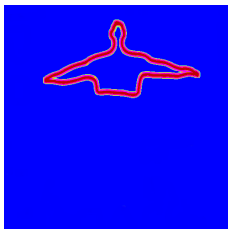


Fig. 8. Result of $\mathbf{I}_{clothes}^{IHI}$.

5.2. Positioning trunk region and arm parts

There are three stages to position the trunk region $(x_T^l, y_T^u; x_T^r, y_T^d)$ and arm parts (ARM^l and ARM^r). First, the initial trunk region is estimated based on the relation among face, trunk and hip defined in Fig. 3(a) as well as $\mathbf{I}_{clothes}^{IHI}$. The shoulders ($V_s^l$ and V_s^r) are also initialized at the first stage. Secondly, a distance-angle plot (or d-θ plot) constructed from $\mathbf{I}_{clothes}^{IHI}$ is used to update the trunk region and fix the shoulder joints. Finally, the arm parts can be immediately computed from the obtained information, and another d-θ plot may also be constructed to estimate the position of forearm parts if necessary. The details of the three stages are now presented as follows.

According to the size ratio among face, trunk and hip defined in Fig. 3(a), the initial of trunk region may be first given by $(x_T^{l0}, y_T^{u0}; x_T^{r0}, y_T^{d0})$. In addition, based on the distribution of $\mathbf{I}_{clothes}^{IHI}$, the summation of $\mathbf{I}_{clothes}^{IHI}$ along each side of the trunk region should have the minimum. Hence, the four sides of the trunk regions may be obtained by the following expressions.

$$x_T^l = \left\{ x \left| \min_{x=x_T^{l0}}^{x_T^{l0}+(x_T^{r0}-x_T^{l0})/2} \left(\sum_{y=y_T^{d0}}^{y_T^{u0}} I_{clothes}^{IHI}(x,y) \right) \right. \right\} \quad (12)$$

$$x_T^r = \left\{ x \left| \min_{x=x_T^{l0}+(x_T^{r0}-x_T^{l0})/2}^{x_T^{r0}} \left(\sum_{y=y_T^{d0}}^{y_T^{u0}} I_{clothes}^{IHI}(x,y) \right) \right. \right\} \quad (13)$$

$$y_T^d = \left\{ y \left| \min_{y=y_T^{d0}}^{y_T^{d0}+(y_T^{u0}-y_T^{d0})/2} \left(\sum_{x=x_T^{l0}}^{x_T^{r0}} I_{clothes}^{IHI}(x,y) \right) \right. \right\} \quad (14)$$

$$y_T^u = \left\{ y \left| \min_{y=y_T^{d0}+(y_T^{u0}-y_T^{d0})/2}^{y_T^{u0}} \left(\sum_{x=x_T^{l0}}^{x_T^{r0}} I_{clothes}^{IHI}(x,y) \right) \right. \right\} \quad (15)$$

The left and right shoulders are initially set to be $V_s^l = (x_T^l, y_T^u; 0, 0)$ and $V_s^r = (x_T^r, y_T^u; 0, 0)$, respectively. Based on the found initial points of (x_T^l, y_T^u) and (x_T^r, y_T^u), we map them onto the image $\mathbf{I}_{clothes}^{HI}$ and construct a useful distance-angle plot as follows.

Assume the activity range of left arm is within $[90°, 270°]$ (the second and third quadrants, left-side of the x-y coordinate system), and that of right arm is within $[270°, 450°$ (or $90°)]$ (the first and fourth quadrants, right-side of the x-y coordinate system). Let (x_T^l, y_T^u) be an origin for the left shoulder, the distance, d, is measured from the origin to the boundary (when a non-clothes pixel encounters) of the clothes region of $\mathbf{I}_{clothes}^{HI}$ along a direction of angle $\theta \in [90°, 270°]$. Thus we may have a half-side d-θ plot in the second and third quadrants for the activity range of left arm. Similarly, let (x_T^r, y_T^u) be an origin for the right shoulder, we may also construct another half-side d-θ plot in the first and fourth quadrants for the activity range of right arm. For the sake of the following analysis, the two half-side d-θ plots are combined together as illustrated in Fig. 9, where $\mathbf{I}_{clothes}^{HI}$ is shown in Fig. 7(a). There are three cases of d-θ plot to be analyzed as follows.

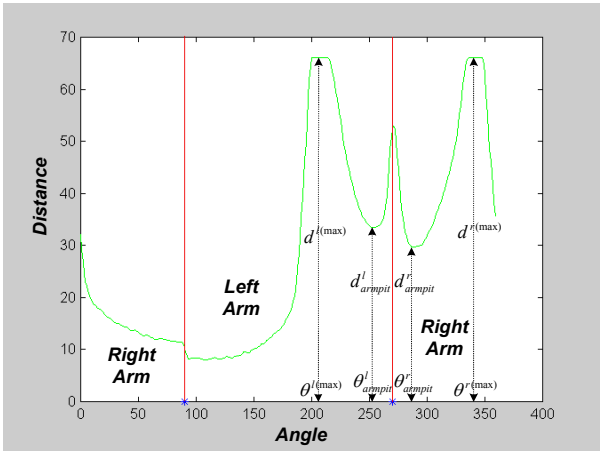


Fig. 9. The d-θ plot for $\mathbf{I}_{clothes}^{HI}$ shown in Fig. 7(a).

(*Case 1*) Fig. 9 shows a general case of d-θ plot for the most of $\mathbf{I}_{clothes}^{HI}$. In such a case, there are two peaks appearing in a half-side d-θ plot, one is near at $270°$, and the other has a maximal distance ($d^{l(\max)}$ or $d^{r(\max)}$ in Fig. 9, for example). Because the trunk region locates between left arm and right arm, we may analyze the d-θ plot starting from angle $270°$ to check whether the known left

(x_T^l) or right (x_T^r) boundary of trunk region is reasonable or not. Now consider the left arm range of the d-θ plot, if there exists a peak nearing at $270°$ on the d-θ plot, then we can find a valley indicated by θ_{armpit}^l, which is the angle of left armpit between left arm and trunk. This indicates that the left boundary of trunk (x_T^l) may be adjusted to left. Only a little amount of space (Δx_T^l) is used to update (x_T^l) for each time, here we define

$$\Delta x_T^l = \frac{1}{2} d_{armpit}^l \sin(\theta_{armpit}^l - 270°). \qquad (16)$$

Similarly consider the right arm range of the d-θ plot, if there exists a peak nearing at $270°$ on the d-θ plot, then we can find a valley indicated by θ_{armpit}^r, which is the angle of right armpit between trunk and right arm. Then Δx_T^r defined as below is used to adjust the right boundary of trunk to right.

$$\Delta x_T^r = \frac{1}{2} d_{armpit}^r \sin(\theta_{armpit}^r - 270°) \qquad (17)$$

For each adjusting time, when a new (x_T^l) or (x_T^r) is adjusted, the down boundary (y_T^d) and the up boundary (y_T^u) of trunk may be obtained by (14) and (15), respectively. Based on the new values, the position of left shoulder joint is updated to

$$(x_s^l, y_s^l) = \begin{cases} \left(x_T^l, y_T^u + \frac{1}{2} L_F\right), & \text{if } 150° \le \theta^{l(\max)} < 270°, \\ \left(x_T^l, y_T^u\right), & \text{if } 120° \le \theta^{l(\max)} < 150°, \\ \left(x_T^l + \frac{1}{2} L_F, y_T^u\right), & \text{if } 90° \le \theta^{l(\max)} < 120°, \end{cases} \qquad (18)$$

and that of right shoulder joint is updated to

$$(x_s^r, y_s^r) = \begin{cases} \left(x_T^r, y_T^u + \frac{1}{2} L_F\right), & \text{if } 270° < \theta^{r(\max)} \le 390° (\text{or } 30°), \\ \left(x_T^r, y_T^u\right), & \text{if } 30° < \theta^{r(\max)} \le 60°, \\ \left(x_T^r - \frac{1}{2} L_F, y_T^u\right), & \text{if } 60° < \theta^{r(\max)} \le 90°. \end{cases} \qquad (19)$$

The offset of $L_F / 2$, defined according to our experiments, is used to make the shoulder information more reasonable on a HB-image. Based on the proposed update rules, for the general case, a reasonable trunk region and shoulder information can be obtained by performing two times of adjustment in our current approach.

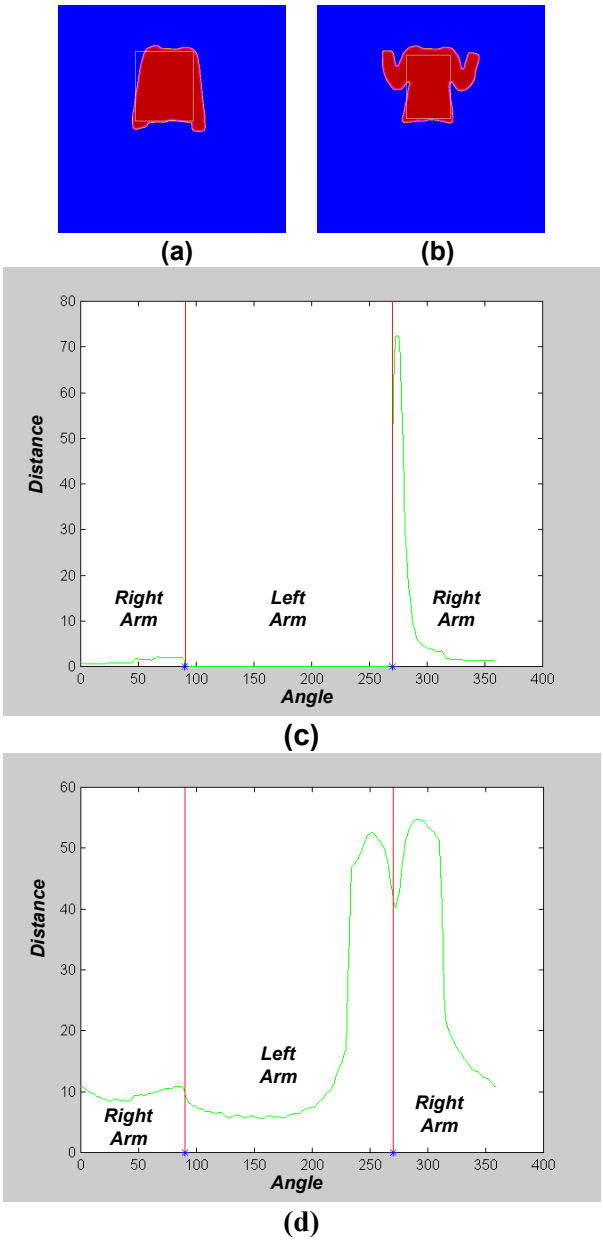


Fig. 10. (a) The right arm of the image $\mathbf{I}_{clothes}^{HI}$ is completely joined with the trunk region, and the initial shoulder of left arm is not located. (b) The arm parts of the image $\mathbf{I}_{clothes}^{HI}$ are clearly distinguished with the trunk region. (c) The corresponding d-θ plot of (a). (d) The corresponding d-θ plot of (b).

(*Case 2*) This case presents only one peak appearing in a half-side of d-θ plot. It happens usually as the following two illustrated examples. Example 1: when the arm part is completely joined with the trunk region, that can be illustrated by the

right arm of the image $\mathbf{I}_{clothes}^{HI}$ in Fig. 10(a), and its half-side d-θ plot in Fig. 10(c). Example 2: when the arm part is clearly distinguished with the trunk region, that can be illustrated by the left and right arms of the image $\mathbf{I}_{clothes}^{HI}$ in Fig. 10(b), and its d-θ plot in Fig. 10(d). The second example can be regarded as a special case of the case 1 presented previously. In the case 2, the boundary of trunk region is not necessary to be adjusted.

(*Case 3*) This case presents zero peak appearing in the d-θ plot as illustrated by the left arm of the image $\mathbf{I}_{clothes}^{HI}$ in Fig. 10(a), and its half-side d-θ plot in Fig. 10(c). When this case occurs, the initial left or right boundary of trunk region should be selected again. The selection can be easily done by moving inward to the left boundary (x_T^l) or the right boundary (x_T^r) along the $(y_T^u)th$ horizontal line until the point (x_T^l, y_T^u) or (x_T^r, y_T^u) is a clothes pixel. Then the d-θ plot is reconstructed to analyze again based on the new values.

After finishing the adjustments of trunk and shoulders, the peaks of the final d-θ plot will be used to estimate the other data of arm parts. Now we present our procedure at this stage using the left arm part. In general, the arm part can be divided into two forms: an extending form (as illustrated in Fig. 7(a)) and a bending form (as illustrated in Fig. 10(b)). In our test, if $d^{l(\max)} > 2.5L_F$, then the left arm belongs to the extending form, otherwise belongs to the bending form.

For the extending form, we may set the length of upper arm and that of forearm to be one half of $d^{l(\max)}$. Thus the results of left arm are completely obtained as follows:

$$\begin{aligned} V_s^l &= (x_s^l, y_s^l; L_s^l, \theta_s^l) \\ &= (x_s^l, y_s^l; \frac{1}{2}d^{l(\max)}, \theta^{l(\max)}), \\ V_e^l &= (x_e^l, y_e^l; L_e^l, \theta_e^l) \\ &= (x_s^l + \Delta x^l, y_s^l + \Delta y^l; \frac{1}{2}d^{l(\max)}, \theta^{l(\max)}), \text{ and} \\ V_w^l &= (x_w^l, y_w^l) = (x_e^l + \Delta x^l, y_e^l + \Delta y^l). \end{aligned}$$

Where $\Delta x^l = \frac{1}{2}d^{l(\max)}\sin(\theta^{l(\max)} - 270°)$ and $\Delta y^l = \frac{1}{2}d^{l(\max)}\cos(\theta^{l(\max)} - 270°)$.

For the bending form, the length of left upper arm may be first determined by $d^{l(\max)} - \Delta d^l$, where Δd^l denotes one half of left upper arm width; and left elbow joint coordinate (x_e^l, y_e^l) can be obtained immediately. Then let (x_e^l, y_e^l) be an origin, we can construct a $\tilde{d} - \tilde{\theta}$ plot (This plot excludes the angle between $\theta^{l(\max)} - 45°$ and $\theta^{l(\max)} + 45°$, since the upper arm has already belonged to this region.) again to find the new peak value $(\tilde{d}^{l(\max)}, \tilde{\theta}^{l(\max)})$, which will be used to determine the left forearm part. Accordingly, we may have the following results for the left arm part.

$$\begin{aligned} V_s^l &= (x_s^l, y_s^l; L_s^l, \theta_s^l) \\ &= (x_s^l, y_s^l; d^{l(\max)} - \Delta d^l, \theta^{l(\max)}), \\ V_e^l &= (x_e^l, y_e^l; L_e^l, \theta_e^l) \\ &= (x_s^l + \Delta x^l, y_s^l + \Delta y^l; \tilde{d}^{l(\max)}, \tilde{\theta}^{l(\max)}), \text{ and} \\ V_w^l &= (x_w^l, y_w^l) = (x_e^l + \Delta\tilde{x}^l, y_e^l + \Delta\tilde{y}^l). \end{aligned}$$

Where $\Delta x^l = (d^{l(\max)} - \Delta d^l)\sin(\theta^{l(\max)} - 270°)$, $\Delta y^l = (d^{l(\max)} - \Delta d^l)\cos(\theta^{l(\max)} - 270°)$, $\Delta\tilde{x}^l$ and $\Delta\tilde{y}^l$ are determined according to the $\tilde{d}^{l(\max)}$ and $\tilde{\theta}^{l(\max)}$. The right arm part may also be obtained with the similar procedure as presented above.

5.3. Positioning hip region and leg parts

The data of face, trunk region and arm parts obtained previously, and the developed procedure, will be used to position the hip region and leg parts. During this stage, the image $\mathbf{I}_{trousers}^{HI}$ shown in Fig. 7(b) will be used for the boundary test of trousers. Now, the hip region is determined first according to the following procedure.

(1) Set $x_H^l = x_T^l$, $x_H^r = x_T^r$, and $y_H^u = y_T^d + 1$.

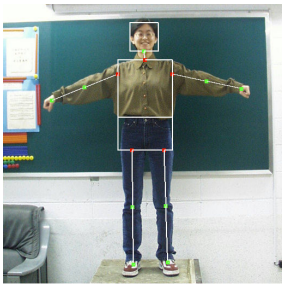
(2) Initially set $y_H^{d(initial)} = y_T^u + 4L_F$ depending on the statistics of person information in our laboratory. In order to find the suitable value of $y_H^{d(final)}$, we define a test block $\mathbf{TB}_{(y^{d(block)})} = (x^{l(block)}, y^{u(block)}; x^{r(block)}, y^{d(block)})$, where

$$\begin{aligned} x^{l(block)} &= x_H^l + \frac{1}{2}\left(x_H^r - x_H^l - L_F\right), \\ y^{u(block)} &= y_H^u, \\ x^{r(block)} &= x_H^r - \frac{1}{2}\left(x_H^r - x_H^l - L_F\right), \text{ and} \\ y^{d(block)} &= y_H^{d(initial)}. \end{aligned}$$

(3) Take the image block $\mathbf{TB}_{(y^{d(block)}+1)}$, compute their (**CCM**)s and (**COV**)s for $\mathbf{TB}_{(y^{d(block)})}$ and $\mathbf{TB}_{(y^{d(block)}+1)}$, then compute the similarity S between $\mathbf{TB}_{(y^{d(block)})}$ and $\mathbf{TB}_{(y^{d(block)}+1)}$.

(4) If $S > 0.8$, then $y^{d(block)} \leftarrow y^{d(block)} + 1$ and go to step 3. Otherwise stop this procedure.

The output of the above procedure will be the final test block $\mathbf{TB}^{final}_{(y^{d(block)})}$ having the desired value $y_H^{d(final)} = y^{d(block)}$.



$(x_T^l, y_T^u; x_T^r, y_T^d) = (210,103;310,208)$
$(x_s^l, y_s^l; L_s^l, \theta_s^l) = (210,129;65,202°)$
$(x_e^l, y_e^l; L_e^l, \theta_e^l) = (149,153;65,202°)$
$(x_w^l, y_w^l) = (89,177)$
$(x_s^r, y_s^r; L_s^r, \theta_s^r) = (310,129;65,348°)$
$(x_e^r, y_e^r; L_e^r, \theta_e^r) = (373,142;65,348°)$
$(x_w^r, y_w^r) = (437,156)$
$(x_H^l, y_H^u; x_H^r, y_H^d) = (210,209;310,268)$
$(x_h^l, y_h^l; L_h^l, \theta_h^l) = (236,268;105,271°)$
$(x_k^l, y_k^l; L_k^l, \theta_k^l) = (237,372;105,271°)$
$(x_a^l, y_a^l) = (239,477)$
$(x_h^r, y_h^r; L_h^r, \theta_h^r) = (295,268;103,272°)$
$(x_k^r, y_k^r; L_k^r, \theta_k^r) = (298,370;103,272°)$
$(x_a^r, y_a^r) = (302,473)$

Fig. 11. The final found HB-framework for the image shown in Fig. 1(a).

Let $(x_{trousers}^{l(boundary)}, y_H^d)$ and $(x_{trousers}^{r(boundary)}, y_H^d)$ be the left boundary and right boundary of the $(y_H^d)th$ horizontal line crossing the image $\mathbf{I}_{trousers}^{HI}$. The hip joint coordinates are then initially set to be $(x_h^l, y_h^l) = (x_{trousers}^{l(boundary)} + L_F/2, y_H^d)$ and $(x_h^r, y_h^r) = (x_{trousers}^{r(boundary)} - L_F/2, y_H^d)$. Also, let $(x_{trousers}^{concave}, y_{trousers}^{concave})$ be a concave point between two legs. We can use the $(x_{trousers}^{concave})th$ vertical line crossing the image $\mathbf{I}_{trousers}^{HI}$ to distinguish the two legs. The function of the $(x_{trousers}^{concave})th$ vertical line is alike to that of "trunk." If it does not exist the concave point, then the $((x_{trousers}^{l(boundary)} + x_{trousers}^{r(boundary)})/2)th$ vertical line will be used instead of the $(x_{trousers}^{concave})th$ vertical line. Therefore, the similar work as developed in Section 5.2 for positioning trunk region and arm parts can completely applied here to determining the leg parts. After deriving all data to construct a HB-framework of the image shown in Fig. 1(a), we obtain the final HB-framework as shown in Fig. 11. Note here that, the leg part can also be divided into an extending form and a bending form. In our test, if $d^{l(\max)} > 3L_F$, the leg belongs to the extending form, otherwise it belongs to the bending form.

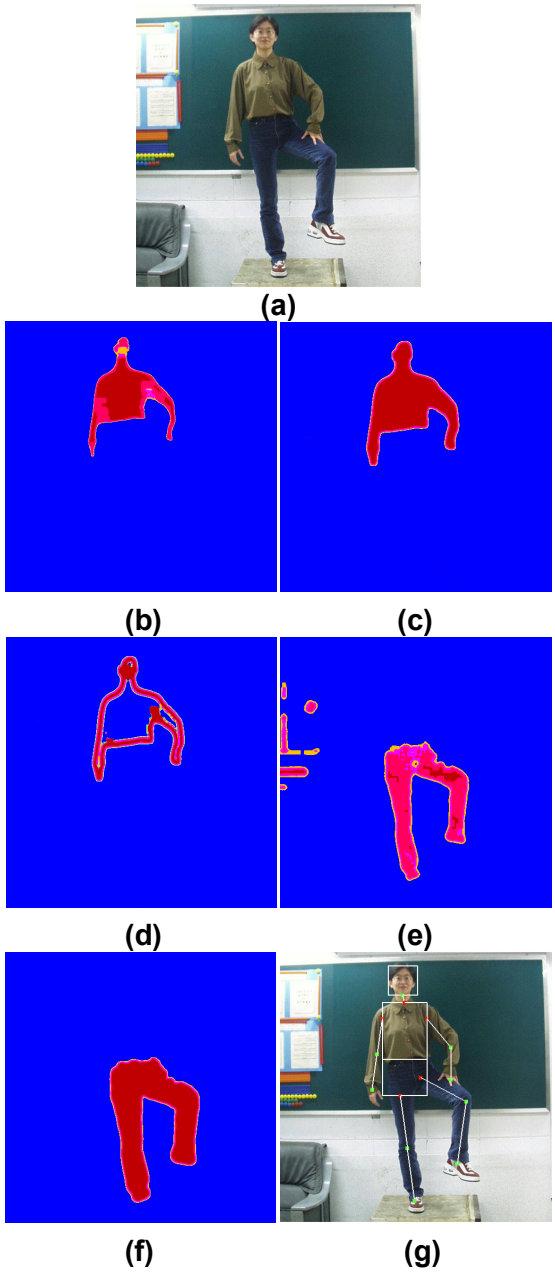


Fig. 12.(a) Original image. Results of (b) $\mathbf{S}_{clothes}$, (c) $\mathbf{I}_{clothes}^{HI}$, (d) $\mathbf{I}_{clothes}^{IHI}$, (e) $\mathbf{S}_{trousers}$, and (f) $\mathbf{I}_{trousers}^{HI}$. (g) The extracted HB-framework shows that the right hip joint is located inside the hip region.

The above example shows a fact that the initial coordinates for both hip joints are located on the hip down boundary, hence the derived leg parts may be reasonable. However, if the initial hip joint coordinate is not located on the hip down boundary, it seems to be necessary to modify the hip joint coordinate. The modification scheme is very simple based on the concept of d-θ plot used in this paper. Take the right leg part for illustration, let the found (x_k^r, y_k^r) be an origin, we can construct a $\tilde{d}-\tilde{\theta}$ plot to find the peak value $(\tilde{d}^{(\max)}, \tilde{\theta}^{(\max)})$ falling into the hip region. Then the new right hip joint (x_h^r, y_h^r) may be readily determined according to the $\tilde{d}^{(\max)} - \Delta d$ and $\tilde{\theta}^{(\max)}$, where Δd denotes one half of right thigh width. Such a case can be illustrated in the example shown in Fig. 12, where (g) shows that the right hip joint is located into the hip region.

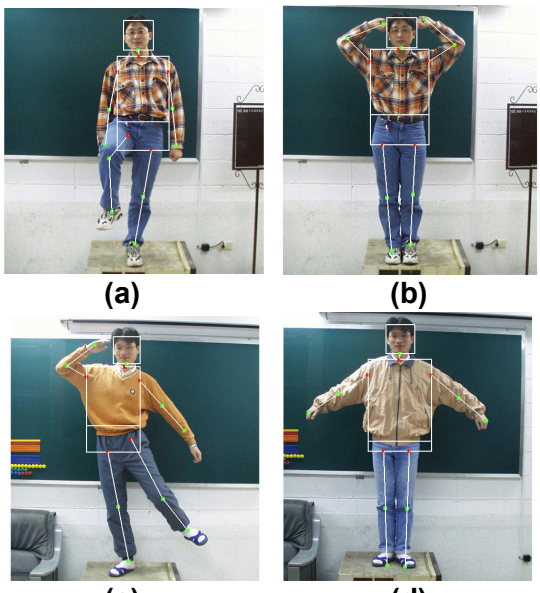


Fig. 13. Some other results of our experiments.

6. HB-framework extraction

Many front-viewed standing HB-images have been experimented with the proposed approach. Persons in our Computer Vision Laboratory of Yuan Ze University were captured with the digital camera, Olympus C-1000L, for experiments. Each color image is of size 512×512 (24 bits/pixel). The proposed algorithms were written in Visual C++5.0 and run on an Intel Pentium III-based personal computer. Fig. 13 shows some other results of our experiments. Although most of extracted HB-frameworks from the front-viewed standing HB-image are reasonable, the results may be affected by the following factors.

- Influence between the color of clothes and that of hand skin.
- Influence between the color of trousers and that of shoes.
- Influence of self-clothes and self-trousers.
- Ratio limitation of the framework with the detected face size.
- Incomplete information of trousers segmentation.
- Unclear trousers boundary occurring between left and right legs.
- Hip down boundary depending on the detected face size.
- Hip joint possibly located inside the hip region.

Based on the above observations from our experiments, it shows the fact that the extraction of HB-framework from a static image is a rather difficult topic at present. Some limits of our approach may be listed as follows.

- The front-viewed face should be at normal top of the trunk.
- The HB-framework is designed to be front-viewed.
- Only the long sleeves of clothes and the long trousers are allowed.
- Some scales xL_F used in the proposed method depend on the statistics of a set of limited human body information.
- The arm parts (and leg parts) cannot overlap each other.

7. Conclusions and future works

In this paper, a basic framework of human body has been defined and an automatic approach for extracting the HB-framework from a given color image has been presented. Our experiments have shown that the main parts of a HB-framework can be extracted reasonably from a given front-viewed standing HB-image by the proposed approach. Some phenomena appearing in our experiments and the inherent limits of the proposed approach have been pointed out. Based on these, A few further works may be continued may include:

- the same color texture problem of clothes and trousers,
- the problem of short sleeves of clothes or of the short trousers,
- the overlapping case or occultation case of arm parts or leg parts,
- the non-front-viewed (e.g. the side-viewed, the back-viewed, etc) and the non-standing HB-image, and
- the case of dressing skirt which hides the leg parts.

In addition, based on the proposed approach, it is possible to extract a 3D HB-framework by using a set of HB-images captured simultaneously from multiple cameras. Then the advanced analyses for human posture and human movements may be readily achieved by using the found stable HB-frameworks. When applying the technique of computer graphics to the HB-framework, it is possible to synthesize a controllable 3D human body in virtual reality applications.

References

[1] V. I. Pavlovic, R. Sharma, and T. S. Huang, "Visual interpretation of hand gestures for human-computer interaction: a review," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 677-695, 1997.

[2] A. D. Wilson and A. F. Bobick, "Parametric hidden markov models for gesture recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 9, pp. 884-900, 1999.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 711-720, 1997.

[4] Y. Adini, E. Moses, and S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 721-732, 1997.

[5] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 743-756, 1997.

[6] I. A. Essa and A. Pentland, "Coding, analysis, interpretation, and recognition of facial expressions," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 757-763, 1997.

[7] J. L. Crowley, J. Coutaz, and F. Bérard, "Things that see," *Communications of the ACM*, Vol. 43, No. 3, pp. 54-64, 2000.

[8] A. Jain, L. Hong, and S. Pankanti, "Biometric identification," *Communications of the ACM*, Vol. 43, No. 2, pp. 90-98, 2000.

[9] A. Pentland, "Looking at people: sensing for ubiquitous and wearable computing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, pp. 107-119, 2000.

[10] A. Pentland, "Perceptual intelligence," *Communications of the ACM*, Vol. 43, No. 3, pp. 35-44, 2000.

[11] I. Haritaoglu, D. Harwood, and L. S. Davis, "Ghost: a human body part labeling system using silhouettes," *IEEE Proc. of 14th IAPR International Conference on Pattern Recognition*, Brisbane, Australia, Vol. I, pp. 77-82, 1998.

[12] S. Iwasawa, K. Ebihara, J. Ohya, and S. Morishima, "Real-time human posture estimation using monocular thermal images," *IEEE Proc. of 3rd International Conference on Automatic Face and Gesture Recognition*, Vol. 2, pp. 492-497, 1998.

[13] T. Harada, T. Mori, Y. Nishida, and T. Sato, "Body parts positions and posture estimation system based on presure distribution image," *IEEE Proc. of International Conference on Robotics and Automation*, Vol. 2, pp. 968-975, 1999.

[14] O. Munkelt, C. Ridder, D. Hansel, and W. Hafner, "A model driven 3D image interpretation system applied to person detection in video images, *IEEE Proc. of 14th IAPR International Conference on Pattern Recognition*, Brisbane, Australia, Vol. I, pp. 70-73, 1998.

[15] Y. Guo, G. Xu, and S. Tsuji, "Understanding human motion patterns," *IEEE Proc. of 12th International Conference on Pattern Recognition*, Vol. 2, pp. 325-329, 1994.

[16] M. K. Leung and Y. H. Yang, "First sight: a human body outline labeling system," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 4, pp. 359-377, 1995.

[17] J. C. Cheng and J. M. F. Moura, "Tracking human walking in dynamic scenes," *IEEE Proc. of 14th International Conference on Image Processing*, Vol. I, pp. 137-140, 1997.

[18] A. Nakazawa, H. Kato, and S. Inokuchi, "Human tracking using distributed vision systems," *IEEE Proc. of 14th IAPR International Conference on Pattern Recognition*, Brisbane, Australia, Vol. I, pp. 593-596, 1998.

[19] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 780-785, 1997.

[20] Y. Li, S. Ma, and H. Lu, "Human posture recognition using multi-scale morphological method and Kalman motion estimation," *IEEE Proc. of 14th IAPR International Conference on Pattern Recognition*, Brisbane, Australia, Vol. I, pp. 175-177, 1998.

[21] T. Mori, Y. Kamisuwa, and H. Mizoguchi, "Action recognition system based on human finder and human tracker," *IEEE Proc. of International Conference on Intelligent Robots and Systems*, Vol. 3, pp. 1334-1341, 1997.

[22] J. O'rourke and N. I. Badler, "Model-based image analysis of human motion using constraint propagation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 6, pp. 522-536, 1980.

[23] J. Y. Zheng and S. Suezaki, "A model based approach in extracting and generating human motion," *IEEE Proc. of 14th IAPR International Conference on Pattern Recognition*, Brisbane, Australia, Vol. 2, pp. 1201-1205, 1998.

[24] Y. Shinagawa, J. I. Nakajima, T. L. Kunii, and K. Hara, "Capturing and analyzing stability of human body motions using video cameras," *Computer Animation '97*, pp. 48-57, 1997.

[25] D. M. Gavrila, "The visual analysis of human movement: a survey," *Computer Vision and Image Understanding*, Vol. 73, No. 1, pp. 82-98, 1999.

[26] Y. S. Chen and Y. T. Yu, "Thinning approach for noisy digital patterns," *Pattern Recognition*, Vol. 29, No. 11, pp. 1847-1862, 1996.

[27] Y. S. Chen, "Hidden deletable pixel detection using vector analysis in parallel thinning to obtain bias-reduced skeletons," *Computer Vision and Image Understanding*, Vol. 71, No. 3, pp. 294-311, 1998.

[28] Y. S. Chen, "Segmentation and association among lines and junctions for a line image," *Pattern Recognition*, Vol. 27, No. 9, pp. 1135-1157, 1994.

[29] H. Wu, Q. Chen, and M. Yachida, "Face detection from color images using a fuzzy pattern matching method," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 6, pp. 557-563, 1999.

[30] An available database at http://peipa.essex.ac.uk/ftp/ipa/pix/faces/manchester.

[31] H. P. Kuo, "Face location and recognition using active shape template," *Master Thesis*, Yuan Ze University, Taiwan, ROC, 1999.

[32] C. C. Chang and L. L. Wang, "Color texture segmentation for clothing in a computer-aided fashion design system," *Image and Vision Computing*, Vol. 14, pp. 685-702, 1996.

[33] Y. F. Li and D. C. Tseng, "Color image segmentation using circular histogram thresholding," *Proc. of IPPR Conference on Computer, Vision, Graphics, and Image Processing*, Taiwan, ROC, pp. 175-183, 1994.

[34] M. J. Carlotto, "Histogram analysis using a scale-space approach," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 1, pp. 121-129, 1987.

[35] Y. M. Lim and S. U. Lee, "On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques," *Pattern Recognition*, Vol. 23, No. 9, pp. 935-952, 1990.

[36] D. C. He, L. Wang and J. Guibert, "Texture discrimination based on an optimal utilization of texture features," *Pattern Recognition*, Vol. 21, No. 2, pp. 141-146, 1987.