

# Body Shape Analysis via Image Processing<sup>\*</sup>

Yang Jinyan<sup>1</sup>, Li Yu<sup>1</sup>, Jiang Tao<sup>1</sup>, Wei Yu<sup>1</sup>, Xu Guanlei<sup>2</sup>

<sup>1</sup> Physical Education Room, Dalian Navy Academy, Dalian 116018, China

<sup>2</sup> Department of Ocean, Dalian Navy Academy, Dalian 116018, China  
{xgl\_86 & xxxgglll0}@163.com

**Abstract** - This paper proposed a new method to analyze the body shape without touching via image processing technique. First, the edges of images are extracted via some edge detection methods. Second, these edge images are analyzed by some measures such as body length and width which can effectively disclose the types of body: thin-long, stout or motile. Experiments show that this proposed method can effectively analyze the body shape with high veracity. This is of much importance to the potential applications in athletics selection.

**Index Terms** - body shape, thin-long, stout, motile, image processing.

## 1. Introduction

In sports and athletics, the body shape plays an important role. In many cases, before the selection of athletes or candidates in sports, the body shape is often analyzed in great details to show the potential ability in sports. Now there have been lots of researches on this field such as [1-10] and so on. For example, in many applications, automatic dress size measurement and virtual try-ons, but also for virtual stunt men in movie productions, virtual models of real persons have to be created that are as detailed as possible and faithfully represent the true body skin surface.

To solve these problems some researchers present a system that is capable of estimating the shape of a human body covered or partially covered by clothes given coarse, noisy, hole riddled or even partial 3D geometry. This is achieved using a statistical model of human body shapes and poses, which is similar to work [1-3]. The other approach works by fitting the statistical model (such as [4]) to the recorded data with an iterative approach, while maintaining that the resulting estimation stays in the space of body shapes spanned by the model. This allows people to estimate the body shape of subjects wearing wide and obstructive apparel. While the generated model is a plausible representation of the subject's body, it is, depending on the clothes, not an exact match but rather a best estimate based on what we can perceive. Even for humans it is difficult to guess the body shape of persons wearing, for example, a long coat. Some biometric measures on the other hand, like height, leg length or arm length, can be calculated relatively accurately though. Some researchers such as Balan and Black [5] have recently presented a system based on the SCAPE model which allows them to estimate the body shape of dressed persons given a number of multi-view images or video sequences. The subjects are allowed to wear arbitrary clothes but have to be captured in a number of different poses

or in a longer animation sequence. Their approach also relies on a color based segmentation of the scans into skin and dressed parts, which is used to apply differently weighted error functions in the segmented regions. In contrast, our method is designed to work without a segmentation and from a single input frame. While our input contains more information than a single multi-view input image, significantly more information about the shape of a person can be extracted from several such multi-view frames when pose and clothing are varied. In the motion capture community several researchers have developed methods to deal with wide clothing. Rosenhahn et al. [6] described a system that allows them to track loosely dressed persons in multi-view video. However, they do require a priori knowledge of both the body geometry and the clothes. Some good tracking results of loosely dressed persons have recently been presented by de Aguiar et al. [7] and Vlasic et al. [8]. Similarly, Starck and Hilton [9] present a system for capturing the performance of actors using multi-view camera systems in studio environments. However, neither paper addresses the underlying body geometry and track only the surface deformation. Balan et al. [10] on the other hand, use a SCAPE based model to track humans. They are thus able to estimate body shape from multi-view video but are restricted to tight-fitting garments.

The common approaches used to recover the (static) shapes are: (1) 3-D scanners (e.g. [1-10]): they are expensive but simple to use and related software are available to edit and model the obtained point clouds. Body scanners usually capture the shape of the entire human body in ca. 20 s. They use the triangulation principle, with laser light or pattern projection method and a CCD camera(s). Their resolution is approximately of 2mm and they can also acquire color texture. The results are precise 3-D data of the subject that are then modeled with reverse engineering software. (2) Silhouette extraction [11,12]: they use different images acquired around a static person and usually fit a pre-defined 3-D human model to the extracted image data. (3) Computer animation software [13-15]: these splines-based packages allow the reconstruction of 3-D models without any measurements and 3-D meshes are created smoothing simple polygonal elements.

Differently, our method in this paper is based on image processing and related techniques [16-18] to obtain the body shapes from static images easily without much processing.

<sup>\*</sup> This work is partially supported by NSFC Grant #61002052 to Xu Guanlei.

## 2. Body Shapes And Features

In this paper, different from the work in [1-15], we mainly focus on the main types of body shape so that we can simplify our analysis. Shown as in Fig.1, there are some body shapes which clearly disclose the body's shapes: thin-long, stout and motile. Here we will analyze the bodies via image processing to obtain their types.

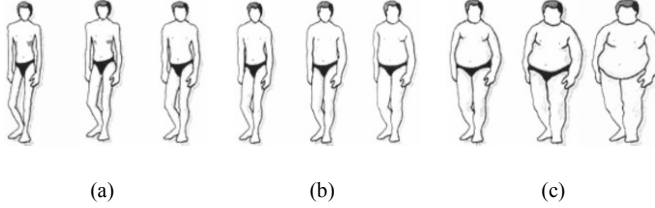


Fig.1. Some body shapes. (a) thin-long ; (b) stout ; (c) motile.

The features of bodies can be defined as four aspects: body length (including stature or height at standing (HS in short), height at seating (HSe in short), the length of arms (LA in short), the length of legs (LL in short) and so on), body width (including width of head (WH in short), width of shoulders (WS in short), width of waist (WW in short), width of antrum (WA in short) and so on), body perimeter (including the perimeter of head (PH in short), perimeter of chest (PC in short), perimeter of arms (PA in short), perimeter of legs (PL in short) and so on), body corporeity (including the substantiation of muscle and so on) which is not used in image processing.

Features of thin-long: thin-long is a relatively concept because if one is thin then he or she will be relatively long. Therefore, thin is the main feature of the thin-long type. Thin-long will mainly include the height at standing, height at seating, the length of arms and the length of legs. In fact, in application, if one is thin-long in whole, or his/her ratio of length and width is high, then he/she will be taken as thin-long.

Features of stout: stout is a relatively concept because if one is thin (fat) and short (long), then he or she will be relatively stout. Stout will mainly include the ratio between height at standing, height at seating, the length of arms, the length of legs and including width of head, width of shoulders, width of waist, width of antrum and so on. In fact, in application, if one is stout in whole, or his/her ratio of length and width is middle, then he/she will be taken as stout.

Features of motile: motile is a relatively concept because if one is thin (fat) and very short, then he or she will be relatively motile. Motile will mainly include the ratio between height at standing, height at seating, the length of arms, the length of legs and including width of head, width of shoulders, width of waist, width of antrum and so on. In fact, in application, if one is motile in whole, or his/her ratio of length and width is low, then he/she will be taken as motile.

In table 1, the main used body measures are listed. These trem's ratios will be employed to analyze the bodies to determine the body shapes. Generally, if the most ratios satisfy

certain conditions, then the body shape will be determined. For example, in table there are 16 measures in whole, and if most of them lie in thin-long scope, then the body will be thin-long, otherwise it will be stout or motile.

Table 1. The body measures used to analyze the body shapes. ("×" denotes this term is used in analysis)

Ratios involved in body measures (left/right)			Body measures											
			body length				body width				body perimeter			
			HS	HSe	LA	LL	WH	WS	WW	WA	PH	PC	PA	PL
Body measures	body length	HS					×	×	×	×	×	×		
		HSe					×	×	×	×	×	×		
		LA											×	
		LL												×
	body width	WH	×	×							×			
		WS	×	×								×		
		WW	×	×										
		WA	×	×										
	body perimeter	PH	×	×			×							
		PC	×	×				×						
		PA			×									
		PL				×								

## 3. Body Shape Measurement via Image Processing

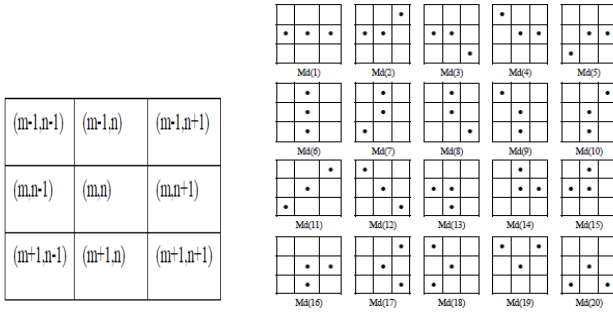
### A. Image segmentation

Image segmentation plays an important role in computer vision and image processing applications [18]. The goal of image segmentation is to produce a separate map of the input image where each uniform textured region is identified with the texture class it belongs to, as is similar with the human's vision system that analyzes and recognizes the different objects with the different textures. With the increasing of the demand of applications in various fields such as medical image analysis and understanding, remote sensing, object-based image coding and image retrieval, image segmentation has received more and more attention over the last several decades and numerous novel methods have been proposed, among many features of which, rotation invariance and robustness to noise and illumination are the few important ones. Early, Chellappa *et al.* used the Gaussian Markov random fields to model texture patterns based on statistical relations between adjacent pixel intensity values [19]. Bovik *et al.* applied the Gabor filters to an image and then computed the average filter responses as features [20]. Mallat introduced the multiresolution wavelet decomposition method, which generates coefficients in the HL, LH, and LL channels for subsequent separate tasks [21]. Weszka *et al.* [22] applied the co-occurrence matrix to extract the mean intensity, contrast, and correlation information from the texture images. Instead, in this paper, we will use the method [22] to segment the images to obtain the bodies.

### B. Edge detection and fill in with black pixels

After the segmentation, we will extract the edges. The output of edge detectors should form continuous boundaries of the regions of interest that can be used for further processing. Edges are image attributes which are useful for image analysis and classification in a wide range of applications. A large number of edge detectors have already been proposed [18] for untextured images. The existing edge detection methods for untextured images can broadly be classified into enhancement/thresholding and edge-fitting edge-detection types. In the case of textured images, these methods cannot be used because these edge detectors are sensitive to micro-edges within textures also. The edge-fitting edge detectors cannot be used because of the absence of any ideal textured edge model. Some methods have already been proposed for finding boundaries between different textured regions. In this paper, we will extract the edges via the approaches as follows.

For image  $f(m, n)$  ( $1 \leq m \leq M$ ,  $1 \leq n \leq N$ ), we assume that  $E_f(m, n)$  is the edge image detected by the edge operator, such as Canny operator and etc.  $EBL(m, n)$  is the  $3 \times 3$  block centered on  $(m, n)$  in edge image  $E_f(m, n)$ . In addition, we assume that the max edge width is not more than 7 pixels.  $Nml(x, y) = 0$  is the normal equation of edge point  $(m, n)$ , and  $Tng(x, y) = 0$  is the tangent equation of edge point  $(m, n)$ , and  $Nml(x, y) \perp Tng(x, y)$ . We assume that all the edge slopes are linear descending or ascending. In Fig.2, twenty edge modes are listed in  $3 \times 3$  blocks. Dots represent the edges in images with value 1. Let the edges be 1 and non-edge pixels be 0. Then the edge mode of is determined by



(a)  $3 \times 3$  image block centered in  $(m, n)$  (b) Edge modes in  $3 \times 3$  block

Fig.2 Edge modes.

$$Md(l) = \arg \min_{l=1,2,\dots,20} \left\{ \sum Md(l) \cap EBL(m, n) = 3 \right\}. \quad (1)$$

According to the directions of normal equation and the estimation of edge parameter, the twenty edge modes can be classified into four types:  $\{Md(l), l=1,2,3,4,5,19,20\}$ ,  $\{Md(l), l=6,7,8,9,10,17,18\}$ ,  $\{Md(l), l=11,15,16\}$  and  $\{Md(l), l=12,13,14\}$ . In the follows, we will give one type of edge parameter estimation.

**Edge parameter estimation of  $\{Md(l), l=12,13,14\}$ .**

$ELD$  is determined by

$$ELD(m, n) = \sup \left\{ |f(m - lp_1, n + lp_1) - f(m - lp_2, n + lp_2)| \right\}, \quad (2)$$

where  $lp_1 = \sup(-3, v_1), \dots, \inf(2, v_2)$ ,  $lp_2 = \sup(-2, v_1), \dots, \inf(3, v_2)$ ,

$$lp_1 \leq lp_2, E_f(m - v_1, n + v_1)E_f(m - v_2, n + v_2) = 1.$$

$ECL$  is determined by

$$ECL(m, n) = \left\{ \sup_l \{f(m - l, n + l)\} + \inf_l \{f(m - l, n + l)\} \right\} / 2, \quad (3)$$

where  $l = \sup(-3, s_1), \dots, \inf(3, s_2)$ ,  $E_f(m - v_1, n + v_1)E_f(m - v_2, n + v_2) = 1$ .

$$EW \text{ is determined by } EW(m, n) = \inf_{u_1, u_2} \{ |u_1 - u_2| \}, \quad (4)$$

where  $u_1, u_2 : \frac{|f(m - u_1, n + u_1) - f(m - u_2, n + u_2)|}{ELD(m, n)} \in [0.9, 1]$ ,

$$u_1 = \sup(-3, s_1), \dots, 0, \quad u_2 = 0, \dots, \inf(3, s_2), \quad E_f(m - v_1, n + v_1)E_f(m - v_2, n + v_2) = 1.$$

$$ECR \text{ is determined by } ECR(m, n) = \frac{ELD(m, n)}{EW(m, n)}. \quad (5)$$

$$ED \text{ is determined by } ED(m, n) = \sup \{ ds_1, ds_2, ds_3, ds_4 \}, \quad (6)$$

where  $ds_1 = \inf_{E_f(u_1, n)=1} \{ |m - u_1| \}$ ;  $ds_2 = \inf_{E_f(m, u_2)=1} \{ |n - u_2| \}$ ;

$$ds_3 = \inf_{E_f(m+u_3, n+u_3)=1} \{ |u_3| \}; \quad ds_4 = \inf_{E_f(m-u_4, n+u_4)=1} \{ |u_4| \}.$$

(6) is the edge distance of single edge point without thinking of connected neighbored edge points. Indeed, connected edge points should be considered as a whole object  $\Omega_{ed}$ . Thus the  $ED$  of a edge connection  $ed(w_m, w_n)$  ( $(w_m, w_n) \in \Omega_{ed}$ ) is defined as

$$ED(\Omega_{ed}) = med \{ ED(w_m, w_n) \}, (w_m, w_n) \in \Omega_{ed}, \quad (7)$$

where “ $med$ ” is the median operator.

Similarly, we can obtain the edge parameter estimation of  $\{Md(l), l=6,7,8,9,10,17,18\}$ ,  $\{Md(l), l=1,2,3,4,5,19,20\}$  and  $\{Md(l), l=11,15,16\}$  in the same manner through different orientations. Note  $ECR$  and  $ED$  are the same for all the four types of edge modes. All the edge parameters are estimated along the normal of edge points. There are some other cases that are not given here. For instance, in a  $3 \times 3$  block there are possibly four or two edge points instead of three ones. The case of four or more edge points has been covered by the twenty edge modes despite of some possible deficiencies.

### C. Measures calculation via pixels and grids

The body measures can be calculated via pixels. After the edge detection, the body measures can be calculated by the following method: the number estimation of grid and pixel.

First, we divide the body's image by the blocks with 8x4 pixels such as shown in fig.3.

Second, find the elements of body: the arms, the legs, the head, the shoulders, the waist, the antrum in the image with coarse image blocks.

Third, count the numbers of the elements of body by pixels in length, we can obtain all the sizes of these body elements.

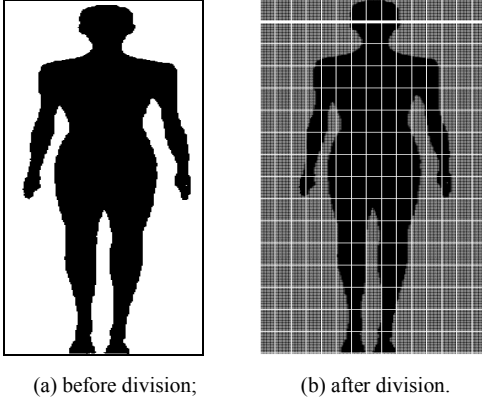


Fig.3. Grid division in the image:

After the three steps, we can calculate the measures of the ratios defined in table2.

Table 2. The correct rate of the body measures estimation via our method.

Correct rate % via our method		Body measures											
		body length				body width				body perimeter			
		HS	HSe	LA	LL	WH	WS	WW	WA	PH	PC	PA	PL
experiments	Experiment 1	82	83	88	83	83	80	83	89	89	87	80	92
	Experiment 2	85	88	86	88	86	82	85	90	81	88	91	87
	Experiment 3	83	90	86	86	89	80	86	81	86	88	87	81
	Experiment 4	83	82	82	88	83	88	89	87	85	87	82	83

### D. Experiments

We test our method with 400 undergraduates of some university in four experiments with every 100 persons being a team. Before test, we indeed know these undergraduates' sport ability through physical training and test. We find our method can effectively give the correct determination of body shapes over 80% in rate.

### 4. Conclusions

This paper we proposed a new method to analyze the body shape without touching the bodies via image processing technique. First, the edges of images are extracted via some edge detection methods. Second, these edge images are analyzed by some measures such as body length and width

which, which can be counted by grids and pixels in images, can effectively disclose the types of body: thin-long, stout or motile. Experiments show that this proposed method can effectively analyze the body shape with high veracity. This is of much importance to the potential applications in athletics selection.

### References

- [1] R. Y. Wang, K. Pulli, J. Popovi'c, Real-time enveloping with rotational regression, in: ACM SIGGRAPH Papers, ACM Press, New York, NY, USA, 2007, p. 73.
- [2] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, J. Davis, Scape: shape completion and animation of people, ACM Trans. on Graphics 24 (3) (2005) 408–416.
- [3] O. Weber, O. Sorkine, Y. Lipman, C. Gotsman, Context-aware skeletal shape deformation, Computer Graphics Forum 26 (3) (2007) 265–274.
- [4] Li Shiming. Research on theory and experiment of bodies' sports via balanced board. Beijing: Beijing Tiyu University, 2003.
- [5] A. O. Balan, M. J. Black, The naked truth: Estimating body shape under clothing, in: D. A. Forsyth, P. H. S. Torr, A. Zisserman (Eds.), Proc. ECCV, Vol. 5303 of Lecture Notes in Computer Science, Springer-Verlag, Marseille, France, 2008, pp. 15–29.
- [6] Xu G., Wang X., Xu X.. Improved bi-dimensional EMD and Hilbert spectrum for the analysis of textures . Pattern Recognition, 2009, 42(5): 718 - 734.
- [7] Yang Jinyan, Li Yu, Xu Guanlei, General Running Pose Analysis via Sparseness, 2nd International Conference on Electric Information and Control Engineering(ICEICE 2012): 4648-4651.
- [8] D. Vlasic, I. Baran, W. Matusik, J. Popovi'c, Articulated mesh animation from multi-view silhouettes, ACM Trans. on Graphics 27 (3) (2008) 1–9.
- [9] J. Starck, A. Hilton, Surface capture for performance-based animation, IEEE Comput. Graph. Appl. 27 (3) (2007) 21–31.
- [10] X. Guanlei, W. Xiaotong, X. Xiaogang. Neighborhood Limited Empirical Mode Decomposition and application in image processing. IEEE Proc, ICIG2007: 149-154.
- [11] A. Hilton, D. Beresfors, T. Gentils, R. Smmith, W. Sun, J. Illingworth, Whole-body modeling of people from multiview images to populate virtual worlds, in: The Visual Computer, vol. 16, Springer-Verlag, Berlin, 2000, pp. 411–436.
- [12] J.Y. Zheng, Acquiring 3D models from sequences of contours, IEEE Trans. Patt. Anal. Mach. Intell. 16 (2) (1994) 163–178.
- [13] D Studio Max 3.1. Available from <http://www.discreet.com/products/3dsmax/> and related plugging <http://max3d.3dluvr.com/main.html> [August 2003].
- [14] Jack. Available from [http://www.eds.com/products/plm/efactory/jack/classic\\_jack.shtml](http://www.eds.com/products/plm/efactory/jack/classic_jack.shtml) [August 2003].
- [15] Maya. Available from [http://www.aliaswavefront.com/en/WhatWeDo/maya/see/solutions/soln\\_intro.shtml](http://www.aliaswavefront.com/en/WhatWeDo/maya/see/solutions/soln_intro.shtml) [August 2003].
- [16] Xu Guanlei, Wang Xiaotong, Xu Xiaogang. Improved Bi-Dimensional Empirical Mode Decomposition based on 2D Assisted Signals: Analysis and Application. IET Image Processing, 2011,5(3):205-221.
- [17] Xu Guanlei, Wang Xiaotong, Xu Xiaogang. On Analysis of Bi-dimensional Component Decomposition via BEMD. Pattern Recognition, 2012,45(4):1617-1626.
- [18] Gonzalez, C.R., Richard Woods, E.: 'Digital image processing' (Pearson Education, 2003, 2nd edn.)
- [19] R. Chellappa and S. Chatterjee, "Classification of textures using Gaussian Markov random fields," IEEE Trans. Acoust., Speech, Signal Process., vol. ASSP-33, no. 4, pp. 959–963, Apr. 1985.
- [20] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, no. 1, pp. 55–73, Jan. 1990.
- [21] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 11, no. 7, pp. 674–693, Jul. 1989.
- [22] J. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of textures measures for terrain classification," IEEE Trans. Syst., Man, Cybern., vol. SMC-6, pp. 269–285, 1976.