Journal of Visual Languages and Computing 47 (2018) 9–18

Contents lists available at ScienceDirect

Journal of Visual Languages and Computing

journal homepage: www.elsevier.com/locate/jvlc

Automatic human body feature extraction and personal size measurement

Tan Xiaohui[a], Peng Xiaoyu[a], Liu Liwen[b], Xia Qing[*,b]

[a] College of Information and Engineering, Capital Normal University, Beijing, China
[b] State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China

ARTICLE INFO

Keywords:
Size measurement
Feature points extraction
Geodesic distance
Heat kernel

ABSTRACT

It is a pervasive problem to automatically obtain the size of a human body without contacting for applications like virtual try-on. In this paper, we propose a novel approach to calculate human body size, such as width of shoulder, girths of bust, hips and waist. First, a depth camera as the 3D model acquisition device is used to get the 3D human body model. Then an automatic extraction method of focal features on 3D human body via random forest regression analysis of geodesic distances is used to extract the predefined feature points and lines. Finally, the individual human body size is calculated according to these feature points and lines. The scale-invariant heat kernel signature is exploited to serve as feature proximity. So our method is insensitive to postures and different shapes of 3D human body. These main advantages of our method lead to robust and accurate feature extraction and size measurement for 3D human bodies in various postures and shapes. The experiment results show that the average error of feature points extraction is 0.0617cm, the average errors of shoulder width and girth are 1.332 cm and 0.7635 cm, respectively. Overall, our algorithm has a better detection effect for 3D human body size, and it is stable with better robustness than existing methods.

1. Introduction

Automatic human body size measurement is very important to personalized products and customization services, especially for garment application [1]. In principle, feature points extraction is always the first essential step, the definition of features should be application-specific, and the accuracy and effectiveness of feature extraction undoubtedly influence the quality of garment application in the most significant way. In the view of ergonomics, during the design or virtual try-on process of human-oriented garment, the human body should be regarded as a core consideration. The 3D virtual garment is constructed to align with the human body feature points or lines. Extracting human feature points or curves can also drive the garment deformations.

Such design automation of human-centered garment products or virtual try-on applications heavily rely on establishing the volumetric parameter of human bodies and feature points usually serves as anchors to constrain this volumetric parameterization. In prior researches, these feature points are always specified by users or semi-automatically selected by rule-based systems [2,3], but how to automatically and properly map certain kinds of focal features from users' psychological/semantic awareness to the right geometric signals is not yet well known.

The automatic extraction on human body is challenging for two reasons. First, the feature points on human body are not always located at the prominent tip location. However, traditional feature points are usually captured on the shape extremities or protrusions, and inevitably are biased towards high shape saliency in some senses, which tends to ignore non-salient but semantically-meaningful regions. Second, the robustness of local shape matching is more problematic when the postures of human bodies are varying or with different body types.

In this paper, we propose an automatic way of size measurement on 3D human body based on extraction of focal features via random forest regression analysis of geodesic distances. The method proposed in this paper aims at providing an automatic way to extract feature points and measure different sizes on 3D human bodies, which usually serves as a pre-processing step for garment designer or virtual fitting application. Feature points can be identified immediately in a universal way. The primary contribution of this paper is a novel modeling approach based on ensemble learning to extract focal features on 3D human body. The predicted displacements of the randomly sampled points on the new 3D human body surface are used to vote for the desired semantic focal features. So the method is automatic in a data-driven way. The scale-invariant heat kernel signature is exploited to serve as feature proximity, so our method is insensitive to postures and different shapes of 3D human body. These main advantages of our method lead to robust and accurate feature extraction and size measurement for 3D human bodies in various postures and shapes.

This paper is organized as follows: Section 2 is an overview of related work. Section 3 is the semantic focal feature definition and the

* Corresponding author.
E-mail address: xiaqing@buaa.edu.cn (X. Qing).

https://doi.org/10.1016/j.jvlc.2018.05.002
Received 18 December 2017; Received in revised form 10 February 2018; Accepted 2 May 2018
Available online 08 May 2018

detailed methodology of our algorithm. In Section 4, the Random Forest regression for local feature identification is presented. And in Section 5, we present some results along with a brief discussion of specific topics and possibilities for the future work.

2. Related work

In practical personalized garment design and virtual fitting application, feature points usually determine the locations of cutting planes and feature lines of the virtual garment. The pre-process of body size measurement is to identify the feature points of the human body, and then simulate the consumer's morphological shape, which provides references for generation of the virtual garment shape [4,5].

Feature detection on 3D shapes, also known as feature extraction, is a classical problem in geometric modeling and computer graphics [6–9]. Researchers intended to focus on the protrusions on 3D shape. Features are often defined in terms of local curvature. All these local geometric features are described by local Gaussian curvature [9], ridge and valley lines [10], prominent tip points [11], the mean curvature flow [10] and the average geodesic distance distribution on the surface [12]. Benefiting from 2D image feature detection methods, some techniques such as 3D-Harris method [13], mesh-SIFT method [14] were extended in the 3D situation. The aforementioned methods depend on the geometric features and the saliency of the points are fully defined by the 3D shape itself while lacking semantic meaning. Shelling points method [15] can combine the geometric properties with users' intention to locate landmark points on new shapes. However, it requires tedious selecting tasks and the landmark detecting process still relied on Gaussian curvature of local geometric properties.

As we can see from the feature extracting process for the non-rigid 3D shape, the methods can essentially boil down to two key components: the descriptor and assignment algorithm. Different approaches such as contour and edge features [16], local patches [17,18], conformal factor [19], differential operators [10], and local volume properties [20] were used as feature descriptors in shape retrieval literature. Unfortunately, none of them satisfies all of the above desired properties such as intrinsic and some of them is not robust to topology, which are important to personalized garment related applications. Dealing with non-rigid 3D shapes requires compensating for the degrees of freedom resulting from deformations. In [21] and follow-up works [22] , authors processed 3D shapes as metric spaces with intrinsic (e.g. geodesic) distances, which are invariant to inelastic deformations. As to 2D shapes, this frame is used with a metric defined by internal distances in 2D shapes [23]. The Laplacian spectra are used as intrinsic shape descriptors [24]. Feature detection based on intrinsic scale-space analysis such as Sochen et al. [25] would find a few reliable points (usually with high curvature), at which scale estimation can be done. In flat regions, no scale estimation is possible. For this reason, In [26], authors avoided feature detection and used a dense feature descriptor computed at every point of the shape in combination with statistical weighting to reduce the influence of trivial points.

The features always represented by grouping regional information in point as the descriptor. The bag of features paradigm relies heavily on the choice of the local feature descriptor that is used with 3D shapes. In non-rigid shape related applications, an ideal feature descriptor should be first of all intrinsic and thus deformation invariant. Second, it should cope with missing parts, and also be insensitive to topological noise and connectivity changes. Third, it should work across different shape representations and formats and be insensitive to sampling. Finally, the descriptor should be scale-invariant. The last two properties are especially important when dealing with shapes coming from different person in virtual fitting application, where shapes appear in a variety of representations and with arbitrary postures. Heat kernel signatures (HKS) is proposed as a deformation-invariant descriptor based on diffusion of multi-scale heat kernels [27]. HKS is a point based signature satisfying all of the good descriptor properties except for scale invariance. The temperature distribution (TD) of the heat mean signature (HMS) as a shape descriptor is defined for shape matching [28]. Their TD is a global shape descriptor and they used L2 norm which is a very basic matching method to compute the distance between two TD descriptors. With a series of transformations, the HKS scale problem is solved [29]. The same research group has proposed Shape Google approach [30] based on the scaled-invariant HKS. The idea is to use HKS at all points of a shape, or alternatively at some shape feature points, to represent the shape by a Bag of Features (BoF) vector. The aforementioned methods are based on intrinsic metric preservation and simply depend on the local/geometric properties.

To overcome the challenge of automatic feature extraction, some machine learning based methods have been proposed which have shown promising performance. Random Forest (RF) based methods are becoming more and more popular. RF was originally proposed by Breiman [31] for general classification or regression, and the class-specific Hough forest was presented in [32] for object detection. Since then, the methods based on RF have shown very promising results in tasks related to landmark detection or organ localization from medical data [33]. The basic idea is to sample local patches (image patches in 2D and volumes in 3D data) in the image, estimate the displacement from the patches to the landmark by RF regression, and then the landmark position is estimated by a voting scheme considering the individual estimations from all the patches. In particular, Criminisi et al. [34], uses regression forest to predict the offsets between the target anatomy and sampled points in the tomography scans and localizes the anatomy's position based on these offsets, and in [31], they used the same idea to automatically detect landmarks in cephalometric X-Ray images.

Inspired by the former work but having sharp contrasts to them, this paper will simultaneously take into account the local information in the neighborhood of points and the symmetrical feature, with an observation that the randomly-sampled points on the shapes can provide convincing vote sources for locating the desired feature points collectively, and give rise to more precise and robust results. At the same time, the efficiency can be enhanced.

3. Semantic focal feature definition

Given 3D human body model is represented as triangle mesh surface $M_S = (V, T)$ with m vertices $V = \{v_1, v_2, ..., v_m\}$, triangle surfaces $T = \{t_1, t_2, ..., t_n\}$ and a set of predefined feature points G_S. We are going to find the corresponding feature points G_H of an input human model H. Without loss of generality, M_H is also represented by a triangle mesh as M_S. The automatic extraction is challenging for two reasons. First, the feature points on human body are not always located in the shape extremities but defined by the requirement of the virtual try-on application, therefore the local shape matching based methods cannot robustly give satisfactory results. Second, the robustness of local shape matching is more problematic when the postures of human bodies are varied (i.e., the 3D bodies are bent). In this paper, we propose an automatic feature point and size measurement method. Firstly, the randomly-sampled points are spread out over the global shape of M_H. Secondly, the final locations of the unknown feature points are collectively decided by the predicted displacements of these randomly-sampled points.

3.1. Garment related sematic focal feature

The critical constraints of proper 3D body and garment measurement are to align 3D garments with body features to guide the later virtual try-on process. The specific points and lines act as defined focal features, which naturally convey the prior knowledge and can be located anywhere on the 3D body shape. In the proposed method in this paper, anatomical landmarks and body features represented by feature points and lines are accurately recognized on both industrial mannequins and non-ideal real-world bodies of various figure shape types.