Human Body Reconstruction

Mingmin ZHANG*[1], Bo WU[1], Kaijia QIU[1], Zhigeng PAN[2]
[1]Zhejiang University, Hangzhou, China; [2]Hangzhou normal university, China

Abstract

This paper gives a survey of different algorithms of human body reconstruction. According to different input data and devices we introduce human body reconstruction algorithms as two classes: image-based methods and depth data-based methods. The first class reconstructs the body with images, which is fast and convenient. However, it is not so accurate as the second one. Reconstruction methods based on 3D depth data differ in the processing of depth data. Some algorithms have accurate results by using total scanning data. Other algorithms estimate human model from part of depth data by using template or fitting in the database.

Keywords: pose recognition, shape reconstruction, 3d body reconstruction

1. Introduction

The traditional method to record and present information of a human body is taking photos or making videos. In the internet age, considering the cost of data storage and network transmission and the reality of the human presentation, most applications still use the media of images or videos. Massive amount of images and videos contains lots of valuable information, especially the human information, which is of significant use in business. Consequently, to obtain the information of images is a main research direction of computer vision.

The researchers analyzed much different information of human bodies, which mainly includes the pose, the shape and the action. The information could be applied to lots of fields. The static images can be searched by the keyword provided by the users based on the content of the images, such as the specific pose or action of humans[1]，or the gender of a person[2,3]. To reconstruct the 3-D model of bodies in the images is of great use. As with the model, the garment can be replaced to achieve the goal of virtual try-on[3,4]. Zhou[6] et al. obtain different dressing results of varies shapes of bodies by reshaping the bodies in the image. The analysis of the dynamic images, i.e. the videos, is used in the vision monitor, animation making and the somatic games. In the environment of high security requirement, such as the bank and the airport, the vision monitor system can judge the danger of a person by estimating the actions of it. Yu[7] et al. estimates the action by classifier, tracking the motion of human bodies and generating the skeleton simultaneously. The vivid animations can be made with a fast speed. The depth image is directly got and used in the somatic games by using the Kinect to easily track the human body.

3D reconstruction technology can be used to acquire and analyze the geometry of the object and environment in the real world. The data obtained can be used to construct digital model of virtual world and also can simplify the model design job. 3D reconstruction is widely used in game design, digital heritage, film making, criminal identification, bioinformatics and reverse engineering[8]. 3D Human information has more value and importance than general scan reconstruction of object and scene in various areas. 3D Human Reconstruction has its unique characteristics: First the body gestures of different person are different but have same characteristics in morphology, thus prior knowledge can be used to reduce the complexity of reconstruction computation and correct the reconstructed models of non-normal body gestures; Second, human cannot maintain static during the scanning process and also in some situation human need to move in real time, which gives the room for template based reconstruction method to be applied. Accurate human model can be acquired by high precise scan device with the subject staying still during a relatively long time scanning, or we must put up with the rough reconstruction results.

In this paper, we will introduce two different types of methods to estimate the pose or reconstruct the human body. The first type is based on the images, which makes of lots of computer vision algorithm. The second type is based on 3D depth data. There are three categories in 3D depth data based reconstruction methods. The key of their difference is how they use the depth data points.

* zhangmm95@zju.edu.cn

2. Image-based pose recognition and shape reconstruction

Large amount of methodology are presented to obtain 3-D information from the 2-D media. A generic method to reconstruct common objects is the Shape From Shading (SFS) technique. Zhang[9] et al. and Durou[10] et al. provided a survey respectively, which are the conclusion and comparison of the SFS technique. However, because of computation complexity, SFS technique is not sufficient to the human body reconstruction. The difficulties or challenges are mainly in five points. 1），The non-rigid body with clothes on it leads to a more complex motion of non-rigid. 2), A series joint on the body makes the degree of freedom high. As a result, fast and valid estimation of high-dimensional parameters is a key problem. 3), 2-D media lacks the details of a dimension and has the problem of self-occluded, which results in the ambiguous and non-unique relationship between 2-D and 3-D. 4), Body pose and motion are under constraints of the physics condition of itself. 5), Complex and changeful pose, background, illumination and garments with the consideration of multi-person, make the human body area extracting difficult.

In the earlier time, the pose is obtained by marking the feature points interactively in the corresponding body image. Lee[11] et al. uses the feature points to estimate pose from single view. They calculate the camera and joints coordinates with these points and with the assuming that relative length of every part is fixed. The fourth problem mentioned above can also be helpful to refine the pose estimated. Taylor[11] et al. also use similar way to get the human pose from an uncalibrated image. Only pose is not sufficient to satisfy some application's requirement, so the researchers construct the 3-D body model by fitting existing model to the image or generating a new one. For instance, Guo[3] et al. employ manual marked feature points similar as described above to estimate the shape of a body.

Getting the pose or shape from images is similar to that from videos. The method can be mainly divided into two classes[13], the discriminative method and the generative method.

There are mainly two kinds of discriminative methods based on the feature used, of which are total match and part match pose estimation. The objects interested are described by features in computer vision. Shape and edge gradient are widely used to extract the human body.

The silhouette is the main feature of shape, which is used by lots of researchers. They extract the silhouette from images by edge detection or image cut approach. GrabCut[14] is a region-based image cut algorithm, which is applied to lots of researches of pose estimation and shape recovery[15-19]. Guan[15] et al. cut the image and get the initial articulated body pose and shape interactively, and then with the SCAPE[20] model's constraint and the body SFS optimization, the result robustly matches the image evidence. SCAPE is a deformable, triangulated mesh model of the human body that accounts for different body shapes, different poses, and non-rigid deformations due to articulation. For vision applications, it offers realism while remaining relatively low dimensional, which is widely used in the reconstruction of body shape[2,3,5,15,21-13]. Balan[21] et al. obtain the pose and shape results through the optimization of a cost function between image observation and a hypothesized mesh. Hasler[19] uses a similar approach with a different model which needs fewer user input compared with SCAPE. Sminchisescu[24] et al. present a novel similarity measure (likelihood) for estimating three-dimensional human pose from image silhouettes and a skeleton-based smoothing method for the image silhouettes that stabilizes and accelerates the search process. In order to make up the information deficiency of 2-D images, some researchers consider to use a multi-view approach to estimate the pose and shape. Hilton[25] et al. extract the silhouettes from views, i.e. the front, back, left and right, with which modify the model and then map the texture in the image to the model to generate animation of the textured model. Lee[26] et al. and Seo[27] et al. did similar research from three views and Lee[27] et al. also made a detailed texture mapping on the face. Seo[28] et al. use the silhouette to retrieve and modify the model by their projection, in which either single or multi-view works. Bǎlan [2] et al. try to estimate the shape under the clothes and infer the gender to modify the shape. Zhu[28] et al. estimate the cross-section of the body by two orthogonal views to reconstruct the 3-D model. Shape context is another feature of the shape, which gives an overall description of a shape by considering the other points distribution histogram to the reference points. Signal[22] et al. and Mori[30] et al. obtain the pose or shape with this feature.

The methods applied to images above are mostly suitable for videos. In some situation, the algorithm needs to be fast, which leads to a simpler model. Peng [31] et al. use the deformation of cylinder to represent the body, which is directly got without the pose estimation. Deutscher [32] et al. apply a simulated annealing particle filter to recover the motion of the articulated body. Fua[33] et al. estimate the pose and state vector with least squares of the silhouette. Theobalt[34] et al. estimate the reflectance and fit the silhouette to the 3-D model database with eight cameras. By texture mapping

from multiple perspectives, they create 3-D renditions of real-world people. Jain[35] et al. realized the tracking and reshaping of the bodies in the video. As the depth camera developing, many researchers start to obtain the body information from depth image. Jiang[5] et al. fit a 3-D morphable human model to the actual body shape of the user in front of the depth camera. Ye[36] et al. present a new view-independent matching algorithm between a 3-D full-body surface mesh and a depth map to estimate body pose configuration from a single depth map. Zheng[23] et al. reconstruct the pose and motion with four synchronized depth images.

The edge gradient feature is common in image researches, which is of high robust. The histograms of oriented gradients (HOG) is employed by many researches, which divides the image into several grid to compute gradient respectively, and then accumulate their directions and finally link to generate the HOG feature vector. Ferrari[16] et al. apply HOG to detect the upper body. Wang[37] et al. combine it with LBP to detect the human bodies with occlusion.

The part-based method finds every part of the body first, such as arms, head, torso, etc., and then judges which kind of pose best fits the combination of the parts obtained. Using the upper-body detector, the body position in Eichner[17] will be cut into head, torso and arms by GrubCut to get the pose finally. Souto[18] et al. apply an artificial neural networks to auto detect the pose with the cut-out body parts. Yang[38] et al. present a general, flexible mixture model for capturing contextual co-occurrence relations between parts, augmenting standard spring models that encode spatial relations. Ferrari[16] et al. develop a generic detector using a weak model of pose and GrubCut to reduce the full pose search space and also an integrated spatial-temporal model to refine pose estimates to estimate 2-D human pose in complex environment. Yu[39] et al. applies action detection and 2-D pose estimation techniques to infer 3-D poses in an unconstrained video. Brauer[40] et al. employ the particle swarm optimization to estimate the 3-D pose with body parts. The silhouette-based method and the part-based method can also be combined to use. Kakadiaris [41] et al. apply the Human Body Part Decomposition Algorithm (HBPDA) to recover all the body parts, and then reconstruct the body model with the silhouette of every parts.

The generative method estimates the pose with a probability optimization model, which combines the prior of every body part and the image information. Chen[4] et al. present a new probability method to estimate the parameters of models to get the shape of body, which can adjust the aspect ratio. Pishchulin[42] propose a new technique to extend an existing training set that allows to explicitly control pose and shape variations. The generative method can also integrate the discriminative method. Sigal[22] et al. employ a discriminative method to get a coarse pose and shape first, and then refine the result with the generative method.

3. Human Body Reconstruction based on 3D depth data

The depth data scanning methods of 3D reconstruction can be divided into two categories named as contact and non-contact methods. The contact-methods acquire data by physical touching with object surface. Such methods can acquire high precision data, and often be used in manufacturing engineering, but it may damage the object to be scanned and is far slower than non-contact methods. Non-contact methods can be divided into passive scan methods and active methods. Passive methods only accept signal passively without emitting energy signal for sensing. Passive scanning methods do not require special hardware. They run on cheap devices and can be used in big scene and object out-door. However passive scanning methods are sensitive to the light and texture of the object and also have high computational complexity and low robustness.

There are so many ways to acquire 3D depth data, they differ in precise and speed. However the key difference of human body reconstruction methods is how to construct human model from scanning depth data. Here we can divide these methods into three categories: registration-based reconstruction methods, database-based reconstruction methods and templates-based reconstruction methods.

Registration-based reconstruction methods construct human model from more than one frame of depth data. They take much time on alignment of all frames, thus have low speed. Templates-based methods always use one frame to finish templates estimation in real time. Database-based methods can use one or more frames, the number of the frames will decide the accuracy of results. Preprocess of scanning depth data is necessary for all three types. We can use mesh-denoise algorithm[43,-45] to mesh of the depth data or directly use point cloud denoise algorithm[46-48] to the raw depth data. Methods of 3D Human Reconstruction have different appearance in accuracy, completeness and computation complexity[49].

3.1. Registration-based reconstruction methods

Registration-based reconstruction methods construct human models by depth data, so they must handle different view of depth data to get all the information of object. Each view of depth data is a depth data frame. The key problem of reconstruction human model from different frames is put points of each frame into one coordinate, which is also called registration. The main steps of registration-based reconstruction methods are preprocess, registration and mesh merge (Fig.1.). Registration in human reconstruction situation is not simple, as people always move a lot. The non-rigid registration algorithms of dynamic objects must be used to solve this problem, which is based on rigid registration algorithm. There are also ways to accelerate the registration based reconstruction methods, since human body has unique characteristics.
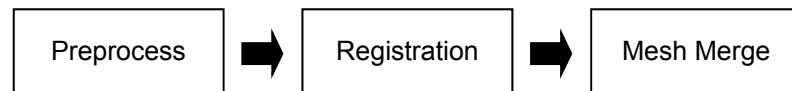

Fig.1. Main steps of registration-based reconstruction methods

The key method to generate model from raw scan data is registration algorithm. To get the total information of object reconstruction method need scanning data of object from different directions. The registration algorithm is used to register different frames of scanning data into one coordinate system. Registration algorithm has two steps: local rigid registration between two frames, global rigid registration of all frames. The classic methods of local rigid registration methods are ICP algorithm (iterative closet point)[50] and RANSAC algorithm (random sample consensus)[51]. ICP algorithm has become the most popular algorithm in 3D Model Registration. This algorithm set the initial registration of two models at first, then set corresponding relationships of points in two models. The corresponding relationships can be used to solve rigid transformation between two models. Using this transformed model as new model to do iterative calculations can get the final registration (Fig.2.).
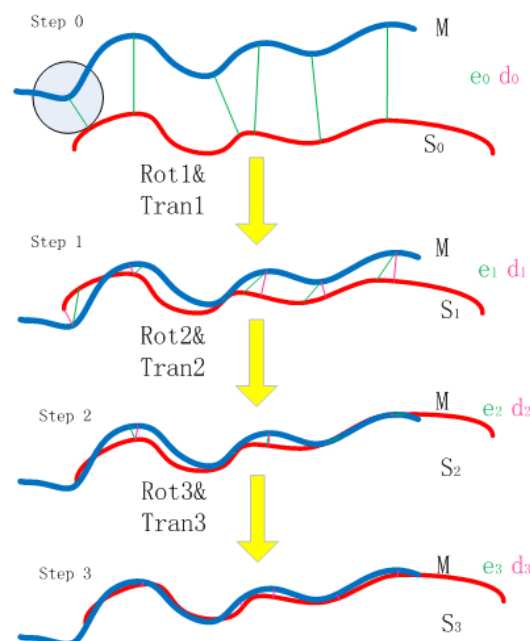

Fig.2. ICP algorithm[52][51]

Local rigid registration algorithm solves the registration problem between two frames of scanning data, but registering all frame one by one using local rigid registration will accumulate the error. Global rigid registration algorithm has solved this kind of problem by distributing error to each frame[53]. Brute-force search method takes the constraints of all frames based on ICP algorithm to solve the equation of all transformations. Brute-force search method cannot be used in practice because of its

large linear equations. Another greedy strategy algorithm registers the new frame to all the registered frames[54], it cannot diffuse the error of registered frames and may not reach the extremism. Based on graph[55,56], global registration method first calculates the local rigid registration of all frame pairs in model geometry level, and then considers the pairs of frames as nodes to construct graph, diffuses global error at a higher level without considering specific geometry of model. Global rigid registration algorithms can minimize the data of all frames acquired from different directions, so it use less memory and can handle models with mass of scanning data.
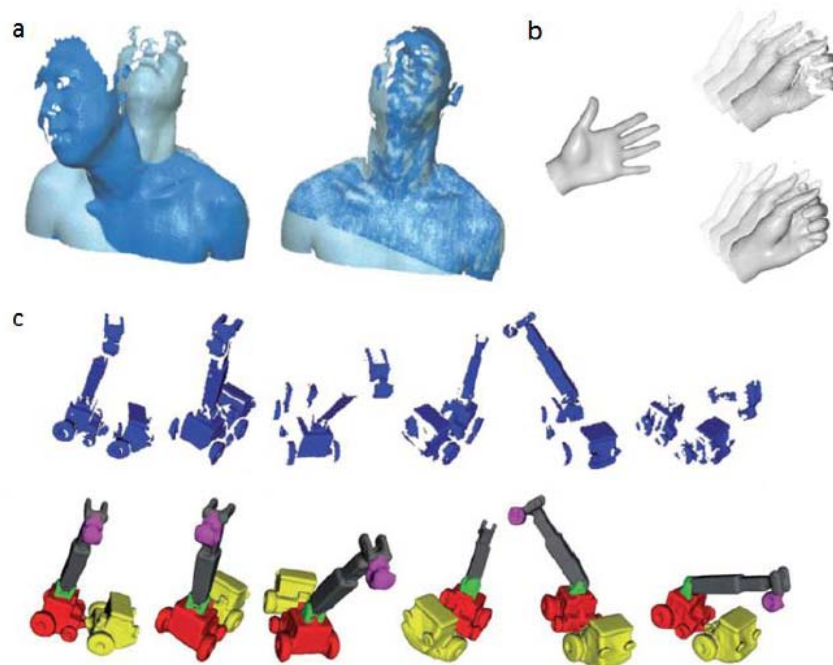

Fig.3. Non-rigid registration algorithm of Dynamic Objects[65]

Some objects are difficult to keep static when the object to be scanned is an animal or Human being. In this situation non-rigid registration algorithms of dynamic objects will be used. Non-rigid registration for dynamic objects is widely studied in recent years[56]. The algorithms can be divided into three types[57]: (a).Non-template non-rigid registration algorithms (Fig.3a.). This type of algorithms has high accuracy of scanning data and low variance of two consecutive scanning frames. Paper[58] makes all the scanning data form a four-dimensional time-space surface, then uses the kinematic characters to track and register numbers of scanning frames. Paper[59] registers two scanning frames in nonlinear deformation by some geometric features and isometric transformation constraints. (b).Non-rigid registration based on accurate template (Fig.3b.). This type of algorithms first need generate a high precision template model, and then use this template to fit all the frames of scanning data. An algorithm based on energy optimization[60] fits template to different frames with the artificial set features, thus to establish a consistent topology body model database. Another algorithm[61] first acquires a more accurate static template model by using 3D scanner, and then fits template to each frame of scanning data by embedded deformation method[62]. (c).Non-rigid registration based on rough template (Fig.3c.). Accurate template grid is often difficult to obtain, but less accurate template is relatively easy to get, thus we can use body skeleton model as rough template to fit joint-animals. Paper[63] first cuts the initial frame data by hand, then accumulates the geometric information which has been captured by identifying and tracking the rigid module of each frame of scanning data. A global registration algorithm of joint-object[64] has been proposed which simultaneously optimizes the registration of multi-frame and skeleton template model.

These three types of methods, as mentioned above, have both advantages and disadvantages. The first type methods require higher input data quality, and pay more time on computation. The second type methods require precise templates, which are difficult to obtain in practical applications, especially for dynamic objects. The third type of methods currently can only deal with a few simple deformations. Research on global non-rigid registration is much less than rigid registration, Occlusion and error accumulation problem in non-rigid registration remain to be further studied[64].

The difficulty of registration algorithm can be reduced by taking advantage of pre-knowledge. The freedom of algorithm can be significantly reduced[61][68] by fitting human model template deformation to each scanning frame (such as deformation embedded algorithm[62]). The human body's movement has its unique characteristics, the motion can generally be thought as rigid motion of segments in global and non-rigid motion in local. The general non-rigid registration can be simplified to a rigid registration of multi-segments. Human body model can be reconstructed by doing global registration of human scanning data first[63] and then local non-rigid registration[64].

Precise reconstructed human model can always be got by using reconstruction methods based on registration. Although there are many ways to accelerate the registration algorithm, but the alignment of numbers of frames is a time consuming job. That's why registration based reconstruction methods cannot work in real time.

3.2. Database-based reconstruction methods

Database-based reconstruction methods use parametric human model database as the pre-knowledge of human body. The database can reduce the computation complexity and correct some error of reconstructed model by pre-knowledge. This type of methods first constructs database by reliable parametric human models, then uses the parameter of input depth data to select the most fitted model from the database. The main steps of database based reconstruction methods can be seen in Fig.4.
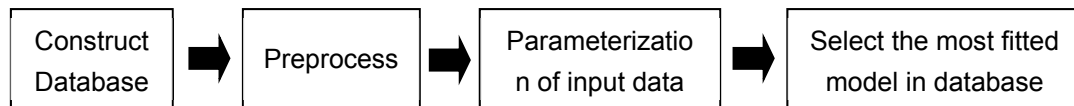

Fig.4. Main steps of database based reconstruction methods

The key problem of database construction is parameterization of human models. A consistent topology 3D Human Body Reconstruction has been proposed[60], which based on human model database. Firstly, the algorithm marked discrete 3D feature points on scanning data of 250 human body, and then register the template model to all the scanning models. Furthermore, PCA algorithm is used to project high dimensional human model database to a lower dimensional space, to build parametric human models. This parametric human model can be applied to Human Data Statistical Analysis, Interpolation of Shape, Human Model Fitting based on discrete feature points. However the method is not applicable when the body postures are different. SCAPE (Shape Completion and Animation of People) [18,19] uses the idea of parametric of human shape to propose a new method. SCAPE considers not only the impact of human body geometry parameters to the human body but also the effects of different body posture. Paper[14] has considered the impact of both body shape parameters and postures parameters to the geometric parameters of human models, in addition, they also set up a European human body database of 550 human models (Fig.5.).
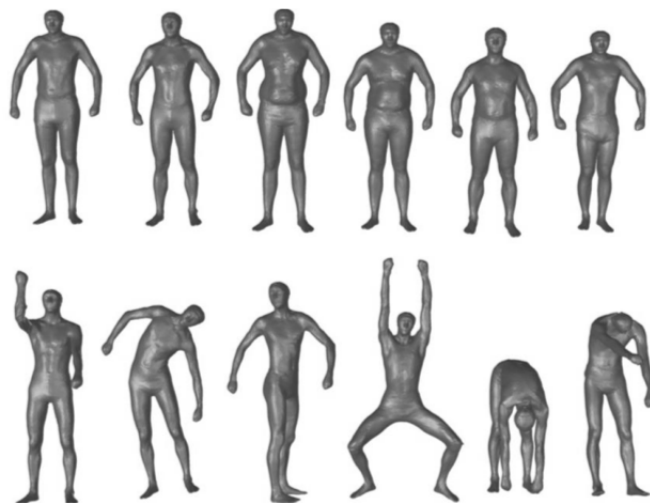

Fig.5. Part of models in database[14]

The computation complexity of human body reconstruction methods is reduced by using priori knowledge of the human body geometry got from the Human Body Database. Using of discrete 3D feature points of human body or body scanning data from single point of view can fit a complete 3D human model based on database[18]. Paper[14] tries to estimate body shape data from color image, by digging the human body part from the image and using silhouette or shadow as constraints to estimate human body shape parameters. Paper[66] reconstructs human body form one single picture by looking for a model from database which is close to human in the image by skeleton and geometry. In the similar opinion, paper[67] restores approximate 3D shape data from body scan data in baggy clothes. In their work the main step to get human body shape data are: rigid registration, morphing of grid, model fitting.

Time and quality of database-based reconstruction methods are depend on the database. The more models in the database the more time will be taken and the more accurate result will be acquired. Database-based reconstruction methods can run in real time only when there are very few models in the database. However we can only get bad result in this situation.

3.3. Templates-based reconstruction methods

Real time reconstruction methods based on database cannot get pleased result. Real time human body reconstruction methods are always based on templates. Templates are parametric human models, which are described by mathematic models, such as GMM (Gaussian Mixture Model). The key idea of templates-based reconstruction methods is estimating parameters of template model from one depth data (3D cloud points). This type of methods can run in real time because it uses one or few frames of depth data and also doesn't have to do registration. Templates-based reconstruction methods first construct templates by parameterizing human model to mathematic models, then use depth data as observed 3D cloud points to estimate mathematic models (Fig.6.).
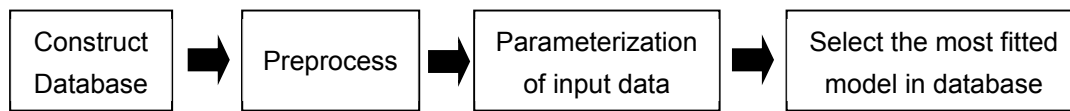

Fig.6. Main steps of database based reconstruction methods

There are always one or more human body templates in Model fitting methods based on templates. One or more accurate human templates are scanned off-line in paper[68][69] which use 3D human body points cloud as input data, use GMM as the parametric model of templates, fit points cloud to templates(Fig.7.), add global and local constraints to templates for getting normal human shape. GMM is an important model to parameterize the human templates[70,71]. Using the number of human template points as the number of GM, the location of points can be the expectation (Fig.8). These methods use precise devices, such as multi-view camera to acquire input data to track articulate objects. The methods about template generation[72,73] can generate template with few interaction without precise template scanning. Such algorithms require high precise input data and templates, so it runs on high price devices and always need to be accelerated by GPU to reach real time calculation. The classic point registration algorithm[74] is used to build relations between of two 3D points sets (Fig.9.).
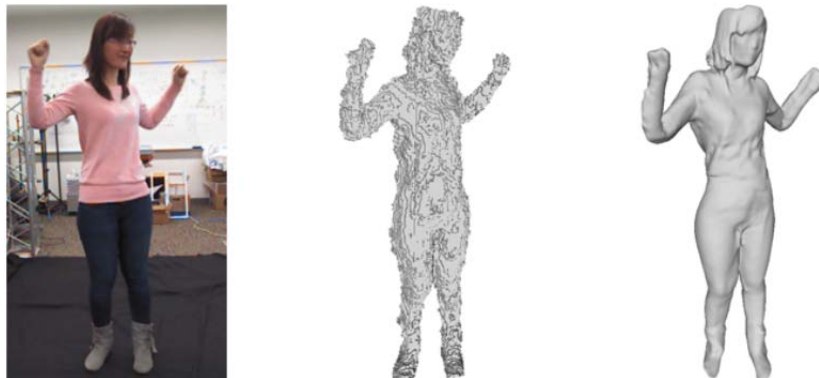

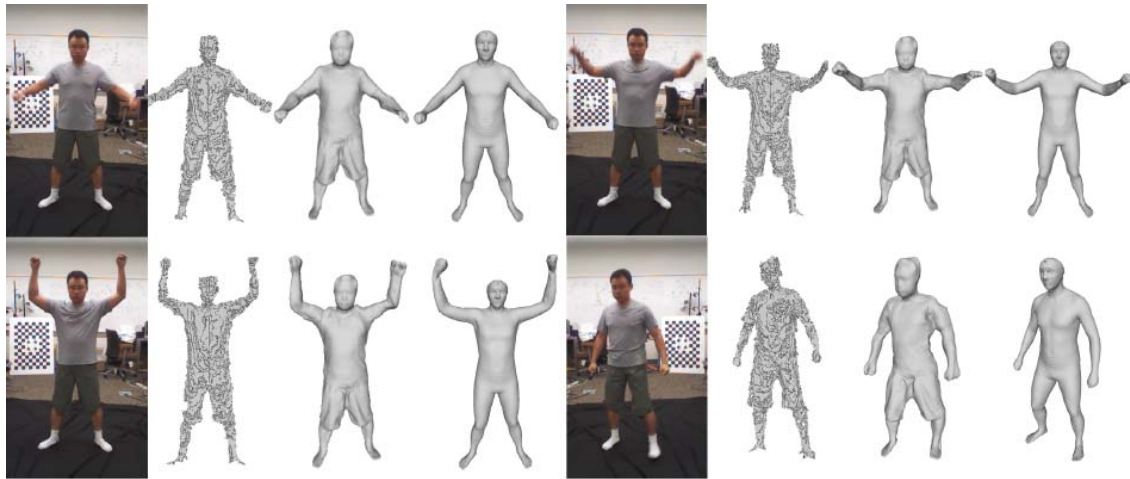*Fig.7. Human Body points fitting algorithm[68]*

Fig.8. Expectation of points location[68]

Fig.9. Points registration algorithm[74]

One template has only one body posture information, thus it will get rough even wrong results when tracking different body postures. To use more templates on real time human body tracking[69,71] is a good idea for getting more accurate results. The key problem of these methods is template selection and registration through different templates. There is also semi-real time algorithm to reconstruct human body model during the natural motion of person[50], which input a series depth data of person's movement to produce a series model movement of constructed human model. Also we can use kinematic chain as template to do point registration and human body reconstruction. Kinematic chain has Kinesiology constraints and can be used to correct unmoral body postures of reconstructed human model[59] (Fig.10.).
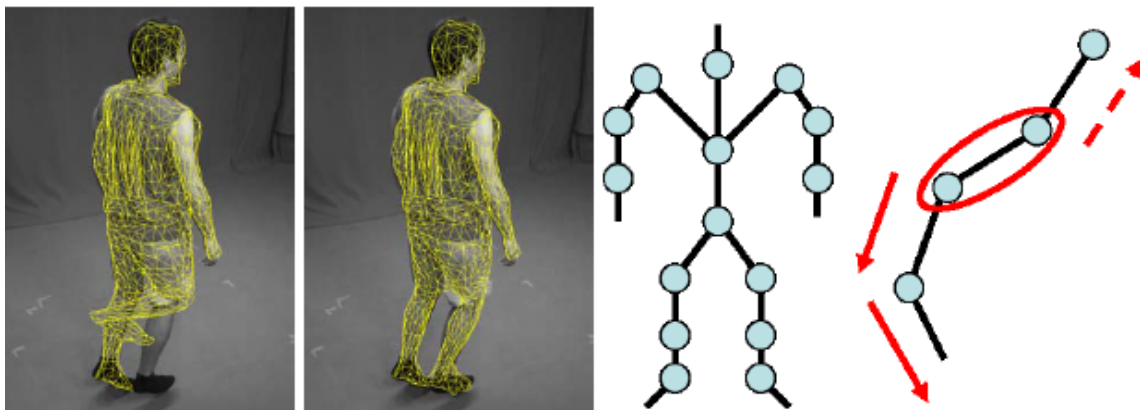

Fig.10. Constraints of kinematic chain template[70]

Although templates-based reconstruction methods can run in real time, but human models are not reconstructed all by raw 3D depth information, thus the precise of model is worse than registration-based reconstruction methods. Templates-based reconstruction methods require high accuracy templates and input data, so it can only runs on high precise real time devices.

4. Conclusion

The human model can be reconstructed from images or depth data. Image-based methods are convenient and do not need data acquirement devices. Human body reconstruction methods based on 3D depth data input data from kinds of scanning devices. The key problem of human body reconstruction methods based on 3D depth data is how to reconstruct human model from depth data frames. And according to that we can divide reconstruction methods based on depth data into 3 categories. The first category, registration-based reconstruction methods can get accurate results, but cannot run in real time. The second category, database-based reconstruction methods, can reduce the computation complexity, but its real time results are very bad. The third category, templates-based reconstruction methods, can run in real time, which needs to implement in high accurate and real time devices (such as 3D camera).

References

[1] *Google Image,* www.google.com, accessed 2015
[2] Bǎlan, Alexandru O., and Michael J. Black. "The naked truth: Estimating body shape under clothing." *Computer Vision–ECCV 2008*. Springer Berlin Heidelberg, 2008. 15-29.
[3] Guo, Yu, et al. "Clothed and naked human shapes estimation from a single image." *Computational Visual Media*. Springer Berlin Heidelberg, 2012. 43-50.
[4] Chen, Yu, Duncan P. Robertson, and Roberto Cipolla. "A Practical System for Modelling Body Shapes from Single View Measurements." *BMVC*. 2011.
[5] Jiang, Bo, Fukai Zhao, and Xinguo Liu. "Observation-oriented silhouette-aware fast full body tracking with Kinect." *Journal of Manufacturing Systems*33.1 (2014): 209-217.
[6] Zhou, Shizhe, et al. "Parametric reshaping of human bodies in images."*ACM Transactions on Graphics (TOG)*. Vol. 29. No. 4. ACM, 2010.
[7] Yu, Tsz-Ho, Tae-Kyun Kim, and Roberto Cipolla. "Unconstrained monocular 3d human pose estimation by action detection and cross-modality regression forest." *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013.
[8] *3D scanner,* http://en.wikipedia.org/wiki/3D_scanner, accessed 2015
[9] Zhang, Ruo, et al. "Shape-from-shading: a survey." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 21.8 (1999): 690-706.
[10] Durou, Jean-Denis, Maurizio Falcone, and Manuela Sagona. "Numerical methods for shape-from-shading: A new survey with benchmarks."*Computer Vision and Image Understanding* 109.1 (2008): 22-43.
[11] Lee, Hsi-Jian, and Zen Chen. "Determination of 3D human body postures from a single view." *Computer Vision, Graphics, and Image Processing* 30.2 (1985): 148-168.
[12] Taylor, Camillo J. "Reconstruction of articulated objects from point correspondences in a single uncalibrated image." *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*. Vol. 1. IEEE, 2000.
[13] Hua, Gang, Ming-Hsuan Yang, and Ying Wu. "Learning to estimate human pose with data driven belief propagation." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 2. IEEE, 2005.
[14] Rother, Carsten, Vladimir Kolmogorov, and Andrew Blake. "Grabcut: Interactive foreground extraction using iterated graph cuts." *ACM Transactions on Graphics (TOG)* 23.3 (2004): 309-314.
[15] Guan, Peng, et al. "Estimating human shape and pose from a single image."*Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009.
[16] Ferrari, Vittorio, Manuel Marin-Jimenez, and Andrew Zisserman. "Progressive search space reduction for human pose estimation."*Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008.
[17] Eichner, Marcin, et al. "2d articulated human pose estimation and retrieval in (almost) unconstrained still images." *International Journal of Computer Vision* 99.2 (2012): 190-214.
[18] Souto Jr, Humberto, and Soraia Raupp Musse. "Automatic detection of 2d human postures based on single images." *Graphics, Patterns and Images (Sibgrapi), 2011 24th SIBGRAPI Conference on*. IEEE, 2011.
[19] Hasler, Nils, et al. "Multilinear pose and body shape estimation of dressed subjects from image sets." *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010.

[20] Anguelov, Dragomir, et al. "SCAPE: shape completion and animation of people." *ACM Transactions on Graphics (TOG)*. Vol. 24. No. 3. ACM, 2005.

[21] Bălan, Alexandru O., et al. "Detailed human shape and pose from images."*Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007.

[22] Sigal, Leonid, Alexandru Balan, and Michael J. Black. "Combined discriminative and generative articulated pose and non-rigid shape estimation." *Advances in neural information processing systems*. 2007.

[23] Zheng, Jiaxiang, et al. "SCAPE-based human performance reconstruction."*Computers & Graphics* 38 (2014): 191-198.

[24] Sminchisescu, Cristian, and Alexandru Telea. "Human pose estimation from silhouettes. a consistent approach using distance level sets." *10th International Conference on Computer Graphics, Visualization and Computer Vision (WSCG'02)*. Vol. 10. 2002.

[25] Hilton, Adrian, et al. "Whole-body modelling of people from multiview images to populate virtual worlds." *The Visual Computer* 16.7 (2000): 411-436.

[26] Lee, WonSook, Jin Gu, and Nadia Magnenat-Thalmann. "Generating animatable 3D virtual humans from photographs." *Computer Graphics Forum*. Vol. 19. No. 3. Blackwell Publishers Ltd, 2000.

[27] Seo, Hyewon, Young In Yeo, and Kwangyun Wohn. "A Data-Driven Shape Model for Human Body Reconstruction from Photos." *Proc. Computer Animation and Social Agents*. 2005.

[28] Seo, Hyewon, Young In Yeo, and Kwangyun Wohn. "3D body reconstruction from photos based on range scan." *Technologies for e-learning and digital entertainment*. Springer Berlin Heidelberg, 2006. 849-860.

[29] Zhu, Shuaiyin, P. Y. Mok, and Y. L. Kwok. "An efficient human model customization method based on orthogonal-view monocular photos."*Computer-Aided Design* 45.11 (2013): 1314-1332.

[30] Mori, Greg, and Jitendra Malik. "Estimating human body configurations using shape context matching." *Computer Vision—ECCV 2002*. Springer Berlin Heidelberg, 2002. 666-680.

[31] Peng, En, and Ling Li. "Acquisition of 3-D surface shape of human body from monocular images without pose estimation." *Control Automation Robotics & Vision (ICARCV), 2012 12th International Conference on*. IEEE, 2012.

[32] Deutscher, Jonathan, and Ian Reid. "Articulated body motion capture by stochastic search." *International Journal of Computer Vision* 61.2 (2005): 185-205.

[33] Fua, Pascal. "Articulated soft objects for multiview shape and motion capture." *IEEE Transactions on Pattern Analysis & Machine Intelligence* 9 (2003): 1182-1187.

[34] Theobalt, Christian, et al. "Seeing people in different light-joint shape, motion, and reflectance capture." *Visualization and Computer Graphics, IEEE Transactions on* 13.4 (2007): 663-674.

[35] Jain, Arjun, et al. "Moviereshape: Tracking and reshaping of humans in videos." *ACM Transactions on Graphics (TOG)*. Vol. 29. No. 6. ACM, 2010.

[36] Ye, Mao, et al. "Accurate 3d pose estimation from a single depth image."*Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011.

[37] Wang, Xiaoyu, Tony X. Han, and Shuicheng Yan. "An HOG-LBP human detector with partial occlusion handling." *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009.

[38] Yang, Yi, and Deva Ramanan. "Articulated pose estimation with flexible mixtures-of-parts." *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011.

[39] Yu, Tsz-Ho, Tae-Kyun Kim, and Roberto Cipolla. "Real-time Action Recognition by Spatiotemporal Semantic and Structural Forests." *BMVC*. Vol. 2. No. 5. 2010.

[40] Brauer, Jürgen, Wolfgang Hübner, and Michael Arens. "Particle swarm optimization on low dimensional pose manifolds for monocular human pose estimation." *SPIE Security+ Defence*. International Society for Optics and Photonics, 2013.

[41] Kakadiaris, Ioannis A., and Dimitri Metaxas. "Three-dimensional human body model acquisition from multiple views." *International Journal of Computer Vision* 30.3 (1998): 191-218.

[42] Pishchulin, Leonid, et al. "Articulated people detection and pose estimation: Reshaping the future." *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012.

[43] Desbrun, Mathieu, et al. "Implicit fairing of irregular meshes using diffusion and curvature flow." *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 1999.

[44] Fleishman, Shachar, Iddo Drori, and Daniel Cohen-Or. "Bilateral mesh denoising." *ACM Transactions on Graphics (TOG)*. Vol. 22. No. 3. ACM, 2003.

[45] Jones, Thouis R., Frédo Durand, and Mathieu Desbrun. "Non-iterative, feature-preserving mesh smoothing." *ACM Transactions on Graphics (TOG)*. Vol. 22. No. 3. ACM, 2003.

[46] Schall, Oliver, Alexander Belyaev, and Hans-Peter Seidel. "Adaptive feature-preserving non-local denoising of static and time-varying range data."*Computer-Aided Design* 40.6 (2008): 701-707.

[47] Fleishman, Shachar, Daniel Cohen-Or, and Cláudio T. Silva. "Robust moving least-squares fitting with sharp features." *ACM Transactions on Graphics (TOG)*. Vol. 24. No. 3. ACM, 2005.

[48] Huang, Hui, et al. "Consolidation of unorganized point clouds for surface reconstruction." *ACM transactions on graphics (TOG)*. Vol. 28. No. 5. ACM, 2009.

[49] Seitz, Steven M., et al. "A comparison and evaluation of multi-view stereo reconstruction algorithms." *Computer vision and pattern recognition, 2006 IEEE Computer Society Conference on*. Vol. 1. IEEE, 2006.

[50] Besl, Paul J., and Neil D. McKay. "Method for registration of 3-D shapes."*Robotics-DL tentative*. International Society for Optics and Photonics, 1992.

[51] Fischler, Martin A., and Robert C. Bolles. "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography." *Communications of the ACM* 24.6 (1981): 381-395.

[52] *ICP*, https://taylorwang.wordpress.com/2012/04/06/iterative-closest-point-algorithm-point-cloudmesh-registration, accessed 2015

[53] Batlle, E., C. Matabosch, and J. Salvi. "Overview of 3d registration techniques including loop minimization for the complete acquisition of large manufactured parts and complex environments." *Eigth International Conference on Quality Control by Artificial Vision*. International Society for Optics and Photonics, 2007.

[54] Masuda, Takeshi, Katsuhiko Sakaue, and Naokazu Yokoya. "Registration and integration of multiple range images for 3-D model construction." *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*. Vol. 1. IEEE, 1996.

[55] Lu, Feng, and Evangelos Milios. "Globally consistent range scan alignment for environment mapping." *Autonomous robots* 4.4 (1997): 333-349.

[56] Sharp, Gregory C., Sang W. Lee, and David K. Wehe. "Multiview registration of 3D scenes by minimizing error between coordinate frames." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26.8 (2004): 1037-1050.

[57] Han Xiaoguang. "Depth Data Based Human Video Editing." *Zhejiang University*. 2011.

[58] West, J. Jason, et al. "Dynamic registration of preablation imaging with a catheter geometry to guide ablation in a Swine model: validation of image integration and assessment of catheter navigation accuracy." *Journal of cardiovascular electrophysiology* 21.1 (2010): 81-87.

[59] Huang, Qi-Xing, et al. "Non-Rigid Registration Under Isometric Deformations." *Computer Graphics Forum*. Vol. 27. No. 5. Blackwell Publishing Ltd, 2008.

[60] Allen, Brett, Brian Curless, and Zoran Popović. "The space of human body shapes: reconstruction and parameterization from range scans." *ACM Transactions on Graphics (TOG)*. Vol. 22. No. 3. ACM, 2003.

[61] Li, Hao, et al. "Robust single-view geometry and motion reconstruction."*ACM Transactions on Graphics (TOG)*. Vol. 28. No. 5. ACM, 2009.

[62] Sumner, Robert W., Johannes Schmid, and Mark Pauly. "Embedded deformation for shape manipulation." *ACM Transactions on Graphics (TOG)*26.3 (2007): 80.

[63] Pekelny, Yuri, and Craig Gotsman. "Articulated object reconstruction and markerless motion capture from depth video." *Computer Graphics Forum*. Vol. 27. No. 2. Blackwell Publishing Ltd, 2008.

[64] Chang, Will, and Matthias Zwicker. "Global registration of dynamic range scans for articulated model reconstruction." *ACM Transactions on Graphics (TOG)* 30.3 (2011): 26.).

[65] Tong Jing. "Scanning and Reconstruction of3D Objects and Human Bodies Using Depth Cameras." *Zhejiang University*. 2012.

[66] Hasler, Nils, et al. "Estimating body shape of dressed humans." *Computers & Graphics* 33.3 (2009): 211-216.

[67] Tong, Jing, et al. "Scanning 3d full human bodies using kinects."*Visualization and Computer Graphics, IEEE Transactions on* 18.4 (2012): 643-650.

[68] Zhang, Qing, et al. "Quality Dynamic Human Body Modeling Using a Single Low-cost Depth Camera." *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on.* IEEE, 2014.

[69] Ye, Mao, and Ruigang Yang. "Real-time simultaneous pose and shape estimation for articulated objects using a single depth camera." *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014.

[70] Gall, Juergen, et al. "Motion capture using joint skeleton tracking and surface estimation." *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009.

[71] Wu, Chenglei, et al. "On-set performance capture of multiple actors with a stereo camera." *ACM Transactions on Graphics (TOG)* 32.6 (2013): 161.

[72] Stoll, Carsten, et al. "Video-based reconstruction of animatable human characters." *ACM Transactions on Graphics (TOG)*. Vol. 29. No. 6. ACM, 2010.

[73] Liao, Miao, et al. "Modeling deformable objects from a single depth camera."*Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009.

[74] Myronenko, Andriy, and Xubo Song. "Point set registration: Coherent point drift." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*32.12 (2010): 2262-2275.