

Making One Structured Body

Won-Sook Lee
Eyematic Interfaces, Inc.
wslee@email.com

Abstract

We discuss how to embed one human body structure on a given human body shape data. The shape data can come from any source, for example designer's tool, or range data scanner (e.g. laser scanner). The main idea is to create approximated body shape starting from generic human by photo-cloning methodology and then make shape-correction using the 3D shape data. In addition, we also describe how to improve the shape such as removing large errors and holes, which come often from range data scanners. Once we have a desired shape, we are able to animate it immediately using generic animation since the animation functionality from the generic model is inherited to the resulted model. In this paper, we described H-Anim body data structure as our desired body structure and experiment several examples from different sources.

1 Introduction

There are various sources to get 3D human data. Compared to face part, the body cloning has been less explored by researchers and is getting more and more popular now. The difference between face and body are first, the body is much bigger and secondly body has skeleton possibly with several skin parts related to the skeleton. Similarly to the face part, there are some methods that concern precision and accuracy and others concerning visual realism. The main methods are classified as (i) laser scanner [5] (ii) silhouette in multiple views on video [7][8] (iii) silhouette in multiple views on photographs [1][3][10][14] (iv) stereo-video [13]. Some of them work very well including both shape and motion capturing, while most of them are suffered by some holes on concave parts on the body.

There are new needs to find one structured shape from different data structures. Homogeneous structure has been an important topic in computer simulation. When we can make one structure from any structure for human shape, the virtual humans become morphable and it makes many things easier, and one of them is to make immediate animation using functional inheritance. The idea is called "generic structure to generic animation". Once shape shares the same structure as the generic structure, the generic animation can be shared too. In this paper, we discuss how to create a structured virtual human with given generic structure, which is standard H-Anim 1.1[6] in our case. We describe not only how to change the structure, but also how to improve the shape itself.

2 Feature-Range based approach

Laser scanners and computer generated modeling using photographs are the most used input methods in practice. Beside the high price or limited accessibility by non professional people, the laser scanner gives a better shape for big regions than computer generated modeling using photographs; meanwhile computer generated modeling using photographs gives better resolution for characteristic regions such as eyes and nose holes. So we propose a methodology not only to make a homogeneous structure of a given shape but also to improve the shape itself by taking

features from computer generated modeling using photographs and taking non-features from scanned data.

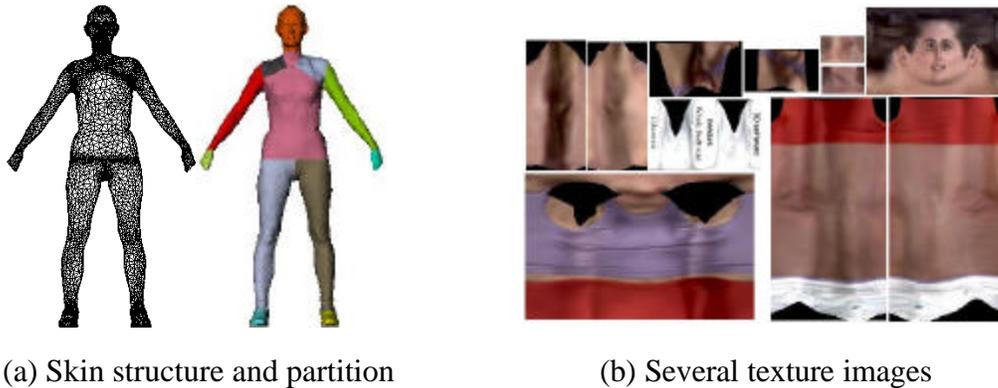


Figure 1: Laser-scanned model has its own skin and texture structure.

Our first example is an unstructured output from laser scanner where 3D shapes and textured information are provided. We collect a model, Tammy.wrz, in VRML 2.0 format from a freeware web site, <http://www.cyberware.com/wb-vrml/Tammy>, which was produced by Cyberware Digitizer™ for human body as shown in Figure 1. The model contains both the face and body parts, where the body is composed of 12 body parts, such as head, larm, lfoot, lhand, lleg, lshoulder, rarm, rfoot, rhand, rleg, rshoulder and torso.

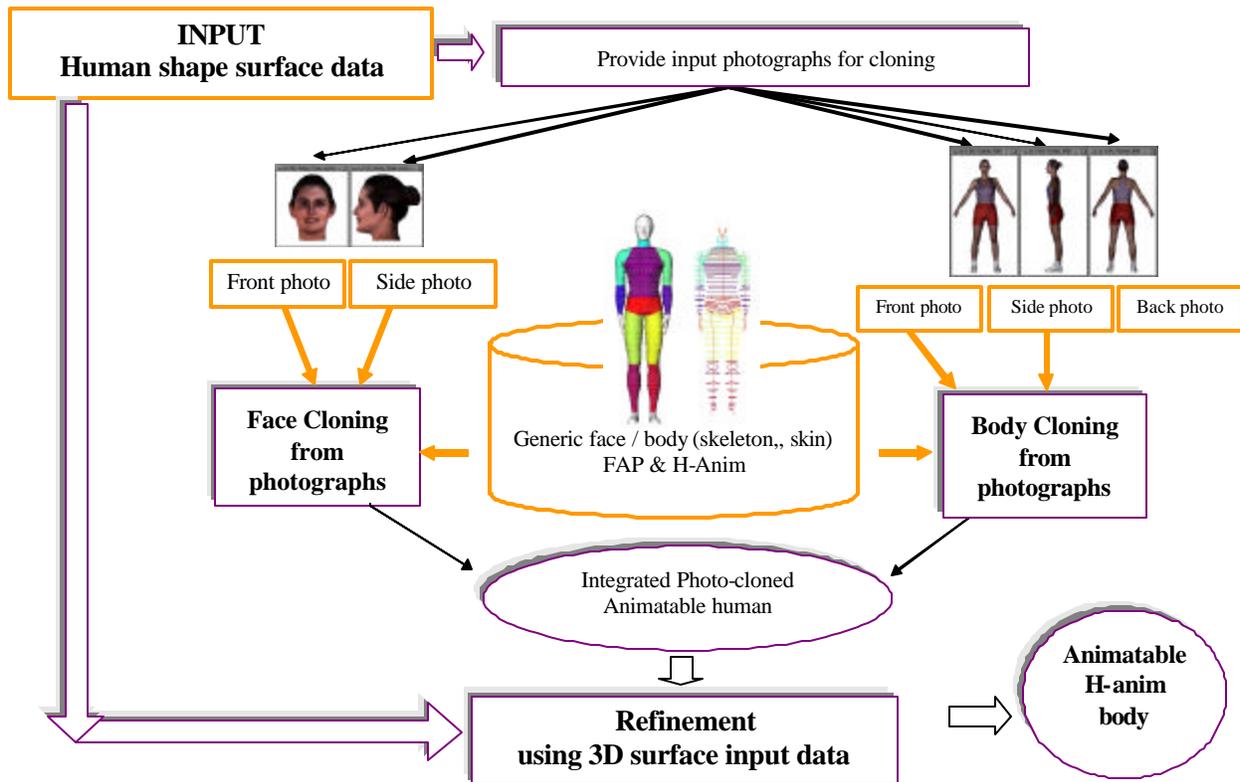


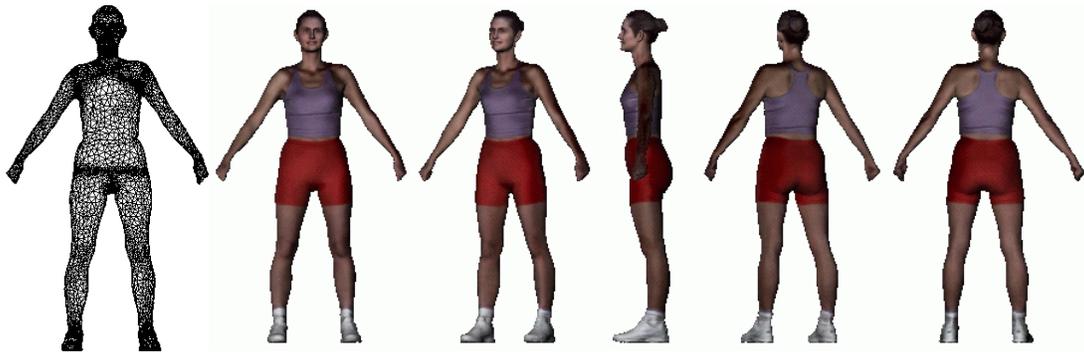
Figure 2: Overflow of face and body cloning.

The overall idea is a combination of feature-based photo-cloning method and surface-based adaptation method as shown in Figure 2.

2.1 Feature-based approach

2.1.1 Photo-cloning

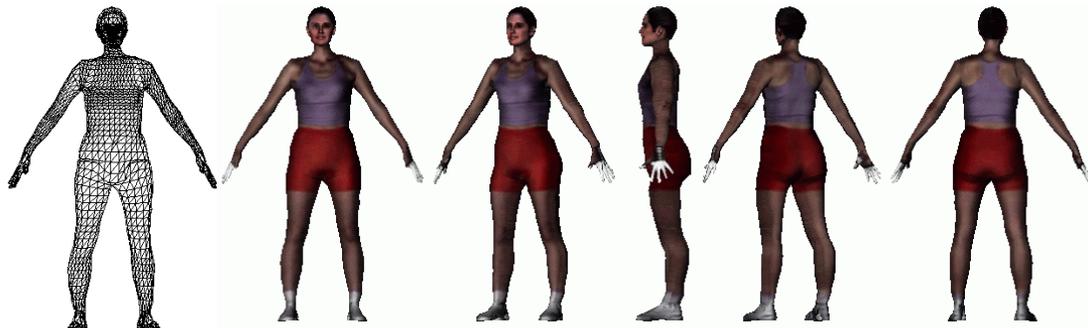
Photographs are one of the easiest and cheapest equipment to get feature points. The features on photographs are very visible most of time under various circumstances. Our previous work [9][10] shows the photo-cloning methods can approximate the person's shape quite well and the texture mapping has a high resolution.



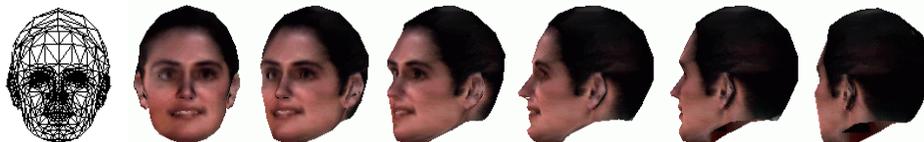
(a) whole body region in several views of the laser-scanned model "Tammy"



(b) face region in several views of the laser-scanned model "Tammy"



(c) whole body region in the several views of the photo-cloned model.



(d) face region in the several views of the photo-cloned model

Figure 3: comparison between the laser-scanned model and the reconstructed model using feature-based photo cloning.

The methodology is composed of two major components: face cloning and body-cloning. The face-cloning component uses feature point detection on the frontal and side images and then uses DFFD for shape modification. Next a fully automatic seamless texture mapping is generated for 360° coloring on a 3D polygonal model. For the body cloning component seeks the solution through a friendly user interface for an accurate localization of feature points with any background, which are used for automatic skeleton modification, automatic edge extraction and for modifying the generic skin into individualized ones, unlike other existing systems that require special environment to obtain input data. A simple but effective heuristics-based boundary extraction algorithm is proposed to automatically extract the body silhouette from the images. Then to avoid the possible overlapping for skin parts, they introduce a two-step modification, first rough matching just with feature points information and then fine matching with detected edge information. The body texture mapping is processed using two images. Moreover, they show how they connect the individualized head to the individualized body to form a complete animatable human model. As a result, they are able to animate the cloned human models in virtual environments and the result can be visualized in any VRML compliant browser.

For our photo cloning, we use snapshot to take five orthogonal photos of the face and body parts after loading the laser-scanned model as shown in Figure 3 (a) and (b).

Five images shown in Figure 2 are used as the input photographs (301x309) for the face and body cloning using the methodology in a grayed region marked as “photo-cloning” in Figure 2. The detailed description of the methodology is found in other literature [9][10].

The animation-structure-embedded virtual human reconstructed from five photographs using feature based approach is found in Figure 3 (c) and (d) which shows the result is very much similar to original range data in visualization while the structure of the skin is completely different from the input data and the skeleton are imbedded inside too. For visual comparison, the snapshots are taken in the same view as the range data in Figure 3 (a) and (b).

However the shape is approximated using only feature points obtained on photographs, there is limitation of the accuracy.

2.2 Surface-based refinement

The result from photo cloning is succeeded to go very near to the shape of the laser-scanned body. The photo-cloned body has reasonably small error so that it becomes easy to remove the gap (error) between the inputted surface model and the cloned model by surface dependent projection. Then the final body results in the same shape from the range data while it has the functional skin structure and skeleton for animation.

However we have to pay attention to the problems of scanned data. The scanned model “Tammy” is a quite good result with smooth surface for face, hair and body without any holes. However for many range data examples from most scanners, the results are not always satisfactory especially for feature regions on a face such as shadowy or complex area around hair, ear, eyes or nose and concave areas on a body such as under the arms or between two legs. Photo-cloned body has the better resolution for the feature parts, for example the eye and lips regions, where range data often have big errors.

So it is desirable to partition the surface depending on feature-superior area and surface-superior area. So we use a methodology to partition the surface into two regions with criteria if they are acceptable or not. We use two methods depending on the range input data. If the range data is as good as ‘Tammy’, we use regional based partition with information in photo-cloned body. For example, we use fine adaptation removing error if the vertex belongs to the region and

if not, we use the photo-cloned vertexes. Of course we perform this measurement after aligning the output surface in the same space of the input surface using scaling and translation. Since there the surfaces have different structure, which means the resulted head has inside structure such as eyes and teeth differently from the laser-scanned head. So we have to make error measurement after taking out the inside structures.

For bodies, we correct the posture to the scanned body. Examples from many scanners shown in Figure 8 are taken with bended arms for a special purpose to make automatic measurement for the scanned bodies, for example, height and upper arms and lower arms are automatically measures inside their system. So this kind of bodies, we have to correct the posture of our H-Anim bodies before processing the find adaptation. Since our bodies cloned from photographs have skeleton and skins attached to skeleton joints, we modify the skin position by modifying skeleton joints. The example used is the scanned body in Figure 8 (a) and (b).

2.2.1 Fine adaptation using error removal

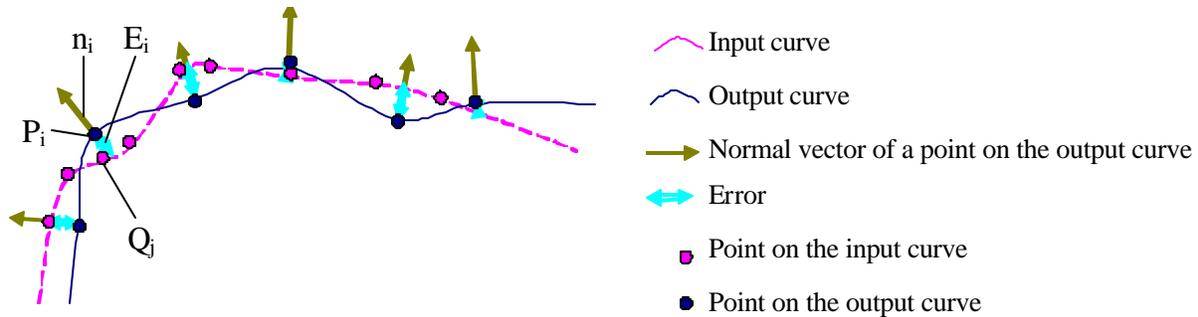


Figure 4: 2D case to show the calculation of the distance for each point on the reconstructed curve to the original input curve by calculating normal vector

Fine adaptation means removing errors between the photo-cloned body and range body for the selected region. So the adaptation depends on what kind of error measurements we are using. In order to compute the error between two surfaces, we use distance measurement between two surfaces. For each point P_i on the output surface, we calculate the normal vector n_i by taking the average of normal vectors of connected triangle faces on the surface. Then we make the projection of n_i to the input surface and get the intersection Q_i . To calculate the projection of n_i , we first calculate the intersection point between the normal vector and each extended plane of a triangle face on the input surface. Then we can decide if the intersection is inside the triangle face or not. Usually there are one or two intersection points of the normal vector on the input surface. We select the nearest intersection point, Q_i , and calculate the distance between the each point P_i on the output surface and the intersection point Q_i . The distance between the point P_i and Q_i is the error E_i of P_i . Figure 4 shows the 2D case.

2.3 Post-processing

2.3.1 Gaps and disordering points

Using error removal process, all points lie on the scanned surface. However it does not guarantee that the surface is smoothed and gap-free and also that the surface points are in proper order keeping the original 'grid' structure, which was designed to make automatic skin deformation and efficient LOD control. For example, see Figure 5 where the first image shows the result from error removing process. The several skin parts have been separated because the

points on boundary between two skin parts have different normal vectors even though they were in the same position. So using the projection method onto the scanned surface, the skin parts became separated. This problem is easy to solve since we have information which skin part is connected to which skin part from the generic model. So after calculating global position, we process sticking points in the same position. The second image in Figure 5 shows the result after sticking. All skin parts are well connected.

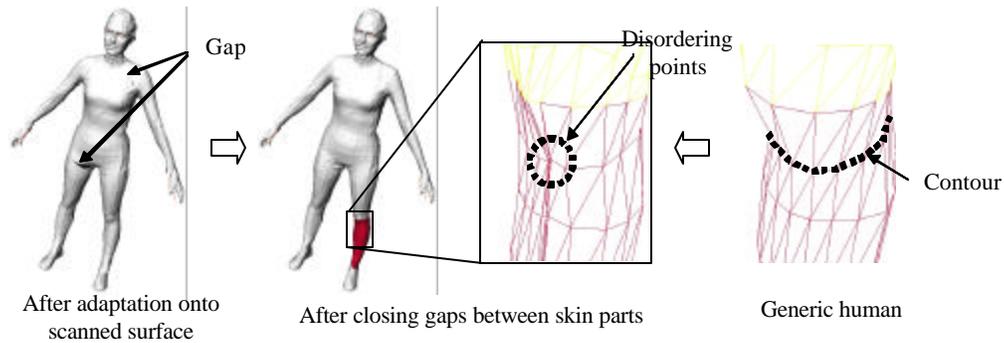


Figure 5: After error removal, skin parts may have gaps and there maybe some disordering for ‘grid’ structure of points.

However there are still problems. As we can see in the magnified image of the second image in Figure 5, there is disordering for ‘grid’ structure of points, where the last image shows the corresponding point structure of the generic model. When there are concave region, the projection method may create some mis-projection. So we correct them by reordering points in the same contour. We select the first point in a contour on a skin part, and reorder points depending on the distance with the fixed orientation either in clockwise or anticlockwise as the same as the generic human. Then if we like to make sure that all points on the same plane, we can have the process to define a plane and then project points on the plane. Then we apply the error removing projection once more. The results are better ordering and fit more to the need of ‘grid’ structure of point’s structure.

2.3.2 Shoulder



Figure 6: Refined photo-cloned body. The shape of the virtual human with the skeleton and skin structure is refined using range data surface.

Shoulder is the most difficult part to get proper adaptation from range data. So we use geometrical correction for shoulder parts using spline interpolation such as Hermite splines. The mid-shoulder point, end-shoulder point and the point in the second next contour on the *upper_arm* after end-shoulder point can be control points. Then the shoulder part has a smoother surface without having big error by wrong projections. Figure 6 is the final result after fine

modification. Compare it with bodies in Figure 3 (a) and (c). It recovered the original inputted shape while photo-cloning method is not able to do it due to lack of feature lines on the region. This method using two-steps modification on the range data is a corresponding method to the “Adaptive mesh” methods by Lee et al. [12] for the face region. Their image analysis technique that searches for salient local minima and maxima in the range image of the subject to adapt the generic face mesh to the laser-scanned range and reflectance data is not always reliable due to the noise. Our two-step feature-range based approach guarantees more robustness and rapid result.

2.3.3 Elbow skeleton correction

Sometimes we encounter different posture for the scanned body, for example the arms are a bit bended. Since our assumption for the arms is straight, the elbow skeleton joint can be far from the appropriate position. So we do skeleton joint correction with given skin envelopes. So we move the elbow skeleton on the center (or it can be positioned with proper bio-mechanical skeleton position with skin envelope) of the skin envelope of elbow, which is a contour in our data structure. However since the skin parts have the corresponding skeleton joints, which transform the local coordinates of the skin part i to global coordinate by 4x4 matrix M_i , we have to adjust the local coordinates of skins of *upper_arm* and *lower_arm* to have the same global position even though the position of elbow is changed. So we calculate the M'_i with updated position of elbow skeleton. Then to get the corrected coordinates P'_i of the coordinates P_i of points on the skin, we apply the following equations.

$$P'_i = P_i M_i (M'_i)^{-1}$$

3 Animation

The photo-cloned human shape inherited the functional structure from generic human, where the points and triangles are designed to be good for the animation. So the resulted body using the proposed method is possible to be animated immediately by applying skeleton and skin animation system. There are several commercial products such as *3D Studio Max* \hat{O} or *Maya* \hat{O} products that support plugins where any mesh can be animated with skeleton simulation by interactive manipulation by skillful designers. However the biggest advantage when we make every body in one kind of skeleton and skin is that we can process generic animation on any individualized body automatically.

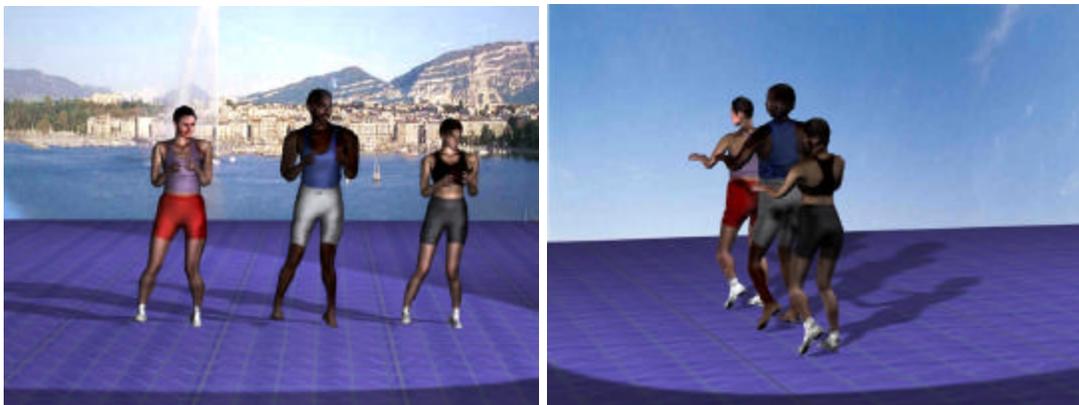


Figure 7: Some examples of animation of three virtual human where the shape is adapted from laser-scanned models.

We use VICON™ motion capture system which is an optical motion capture system to get a skeleton animation and *3D StudioMax* \hat{O} for skin deformation using attachment data, which is a weight assignment for every point on a skin surface on certain skeleton joints.

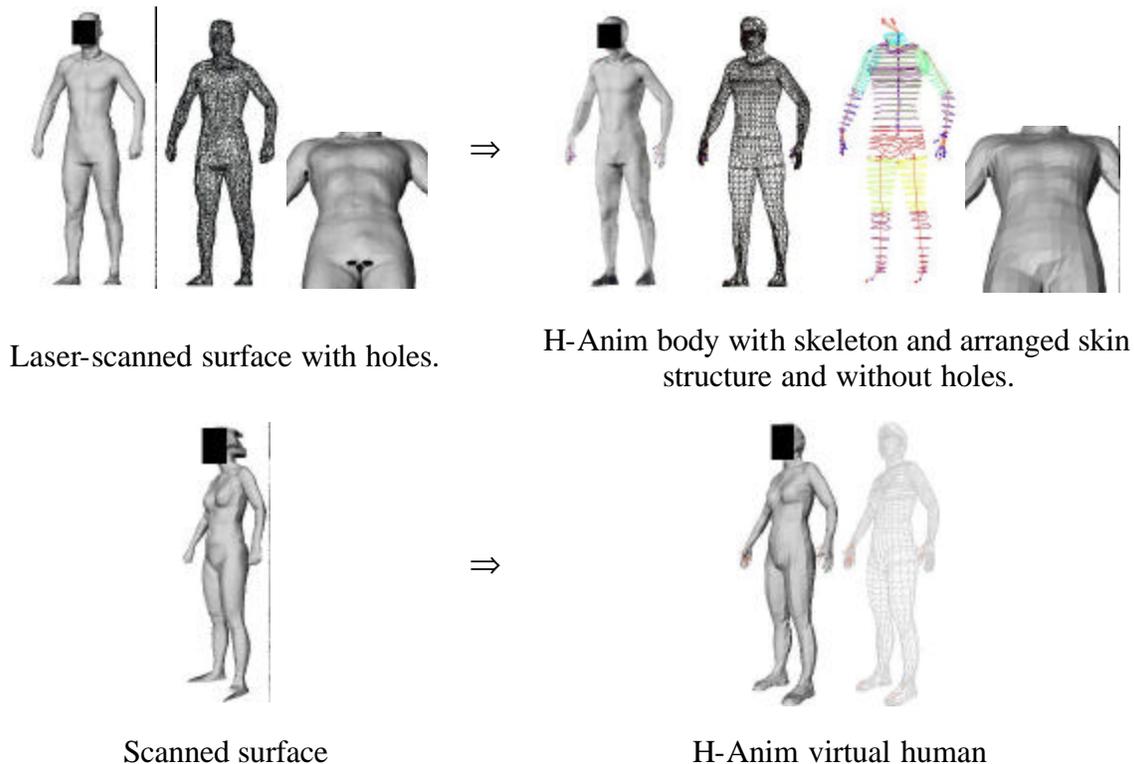
For the motion capture, we have chosen a dancing motion, which is typically non-trivial movement to synthesize. Six cameras whose sample rate is 120Hz are used along with 25mm markers for VICON™ motion capture system.

We collected three models from a website [5] for the laser-scanned human models. They are called as *Tammy*, *Alison* and *John*. *John.wrz*, has the same skin partition with *Tammy.wrz*, but the number of points for the same body part is different while *Alison.wrz*, found in the same website, has different skin partitions such as only 9 skin body parts like larm, lfoot, lhand, lleg, rarm, rfoot, rhand, rleg, and torso.

Even though those three models collected from a web have different structures, we convert them to be the generic structure and apply generic animation to make immediate animation on them. Some snapshots of the three dancing virtual humans are shown in Figure 7.

4 Examples from other source equipment

There are many various sources to get a range data. We collected several surface data from designer's tool such as Curious Labs™ Poser 4 [4], from an infrared scanner, Hamamatsu™ and from another laser scanner, Techmath™. We convert them to the same structured H-Anim bodies.



Laser-scanned surface with holes.

H-Anim body with skeleton and arranged skin structure and without holes.

Scanned surface

H-Anim virtual human

Figure 8: Two laser scanned examples converted to H-Anim virtual humans

See Figure 8 where very high number of points is produced and the detailed body shape is well captured. However the above example surface in Figure 8 shows there are holes on armpits and between legs, which are very common from any laser scanner unless there is a special post-

processing. Below example in Figure 8 shows another problem for hair and feet area. The scanned surface has shape problems on hair and foot even though the number of polygons is 112,629. The hair problems come from the black hair color and the capturing did not work properly. Our several steps process to produce H-Anim body removed all shape problems while also reducing the number of polygons greatly. All bodies shown in this paper have 4746 polygons for the body, hands and feet and 2314 polygons for the face and also they have all information for facial and body animation.

We also process a 3D virtual human model from Curious Labs™ Poser 4 that is a 3D-character animation and design tool for artists and animators. The surface model produced from the system is shown in Figure 9 (a) while the corresponding H-Anim body is shown in Figure 9 (b). Even though the surface structures from difference sources are different, we succeed to produce the same structure, which is H-Anim body. Another source we testes is Infrared scanned body shown in Figure 9 (c) where a female body is scanned. The conversion to H-Anim structure is processed and the result is shown in Figure 9 (d).

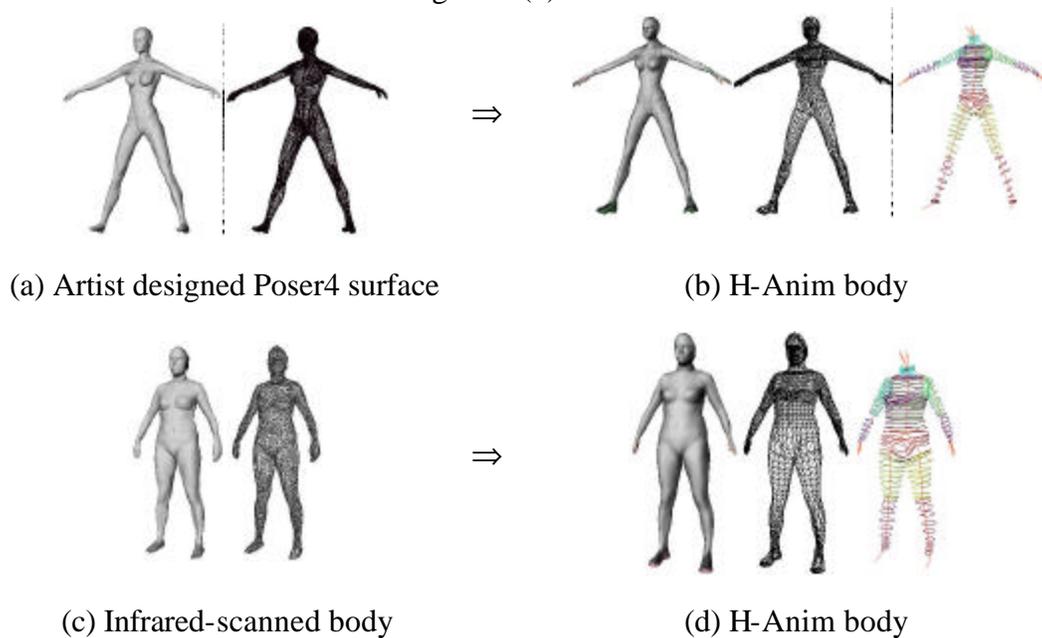


Figure 9: (a)(b) Designed body to H-Anim body. (c)(d) Infrared-scanned body from Hamamatsu \hat{O}

5 Conclusion

The idea to convert any mesh to a given skin structure is useful to solve a problem to make database of morphable human models [11] and to apply immediate animation without going through case-by-case animation functionality embedding.

We have described the methodology and algorithms for creating one structured animatable virtual human from laser-scanned body. We have developed a face-adaptation system and body-adaptation system with a given generic body in H-Anim standard. The result is H-Anim body whose shape is borrowed from scanned shape while the structure from H-Anim structure. In addition since our mythology combines features from computer generated model using photographs and non-features from range data, the final results has improved in terms of shape too.

Future work is an extension of “generic structure to generic animation” idea on more various applications.

6 Acknowledgement

The authors would like to thank members of MIRALab, University of Geneva, for their various kinds of help, especially to Professor Nadia Magnenat-Thalmann, Nedjma Cadi-Yazli, Hyewon Seo, Dimitrios Protopsaltou and Christiane Luible.

7 Reference

- [1] Gu J., Chang T., Gopalsamy S., and Shen H., “A 3D Reconstruction System for Human Body Modeling”, In *Modelling and Motion Capture Techniques for Virtual Environments (Proc. CAPTECH'98)*, (Springer LNAI LNCS Press), pp. 229-241, 1998.
- [2] Guenter B., Grimm C., Wood D., Malvar H., Pighin F., “Making Faces”, *Computer Graphics (Proc. SIGGRAPH'98)*, ACM Press, pp. 55-66, 1998.
- [3] Hilton A., Beresford D., Gentils T., Smith R. and Sun W., “Virtual People: Capturing human models to populate virtual worlds”. In *Computer Animation (Proc. Computer Animation'99)*, IEEE Computer Society Press, pp. 174-185, 1999.
- [4] <http://www.academic-softwares.com/curiouslabs.asp>
- [5] <http://www.cyberware.com/wb-vrml/>
- [6] <http://www.H-Anim.org>
- [7] Kakadiaris I. A. and Metaxas D., “3D Human Body Acquisition from Multiple views”, in *Proc. of the Fifth ICCV*, IEEE Computer Society Press, pp. 618-623, 1995.
- [8] Kakadiaris I. A. and Metaxas D., “Model-based estimation of 3D human motion with occlusion based on active multi-viewpoint selection”, In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society Press, pp. 81-87, San Francisco, CA, June 18-20 1996.
- [9] Lee W., Magnenat-Thalmann N., “Fast Head Modeling for Animation”, *Journal Image and Vision Computing*, Volume 18, Number 4, pp.355-364, Elsevier Science, 1 March, 2000.
- [10] Lee W., Gu J., Magnenat-Thalmann N., “Generating Animatable 3D Virtual Humans from Photographs”, accepted in *Eurographics 2000*, Volume 19, Number 3, *Computer Graphics Forum*, Blackwell publisher, 2000.
- [11] Lee W., Magnenat-Thalmann N., “Virtual Body Morphing”, in *Proc. Computer Animation 2001*, Seoul, Korea, November 6–8, 2001.
- [12] Lee Y., Terzopoulos D., and Waters K., “Realistic Modeling for Facial Animation”, In *Computer Graphics (Proc. SIGGRAPH)*, ACM Press, pp. 55-62, 1995.
- [13] Plänklers R., Fua P., D'Apuzzo N., “Automated Body Modeling from Video Sequences” in *Proc. IEEE International Workshop on Modelling People (mPeople)*, IEEE Computer Society Press, Corfu, Greece, September, 1999
- [14] Weik S., Wingbermühle J., Niem W., “Automatic Creation of Flexible Antropomorphic Models for 3D Videoconferencing”, *Proceedings of Computer Graphics International CGI*, IEEE Computer Society Press, June 22-26, Hanover, Germany, 1998