Individualising Human Animation Models

Xiangyang Ju and J. Paul Siebert

3D-MATIC Laboratory, Department of Computing Science, University of Glasgow, Glasgow, UK
xju@dcs.gla.ac.uk psiebert@dcs.gla.ac.uk

Abstract

Realistic human body animation requires accurate body geometry, realistic textures and believable behaviours. Existing human animation models provide excellent tools to control articulated motion and body surface deformations according to body postures, but modelling a specific individual requires a large degree of skill to sculpt a model resembling that individual. The advent of 3D imaging collection techniques means that the highly accurate 3D surface of a specific person can now be collected in a matter of milliseconds, but most captured 3D surfaces have little semantic information. Consequently, a significant degree of laborious manual intervention is still required to produce a virtual character from their 3D image using traditional animation approaches.

The authors present a method for combining 3D images of real people with existing non-specific, i.e. generic, human animation models to achieve highly realistic human animation models of the 3D imaged individuals. This has been achieved by developing a two-step method for conforming a generic human animation model to fit or "clone" the 3D geometry captured from a specific individual.

A segmentation procedure is first applied to the 3D imaged human body surfaces to impose a semantic labelling of the data into body surface parts. Having thereby established correspondence between the segments of the scanned human body and those of the generic model, a rigid body transformation followed by a global 3D mapping is applied to bring the generic mesh into approximate alignment with the 3D imaged model. The second step comprises a conformation procedure driven by (minimising) the distance between the closest-points between the respective meshes to bring the generic model surfaces into close alignment with those of the real-world 3D surfaces.

1. Introduction

Human body animation is one of the most challenging tasks encountered by animators, especially when creating animation models that represent faithful facsimiles of specific individuals. Various human body animation packages are available commercially to provide useful virtual animation models with articulated geometry structures, controllable human body movements and surface deformation mechanisms. Despite the availability of these animation tools the task of "individualising" synthetic generic models, i.e. tailoring the generic model to the appearance of a specific individual, continues to require a large degree of skill and manual intervention. Various techniques such as laser scanning, Moiré metrology and stereo photogrammetry can recover the realistic 3D body shape of individuals. The highly accurate 3D surface of a specific individual can now be collected in a matter of milliseconds using photogrammetry-based 3D imaging techniques [1], e.g. C3D® technology, originally developed by the Turing Institute and utilised at Glasgow University. The authors present a method for combining 3D images of real people with existing generic human animation models provided by animation package manufacturers to achieve highly realistic human animation models of the 3D imaged individuals. This has been achieved by developing a method for conforming a generic human animation model to fit or "clone" the 3D geometry captured from a specific individual.

Previous investigations into methods for reducing the burden of manual intervention required when creating individual virtual characters include Hilton and Gentils model-based approach for body model reconstruction from a set of low-cost colour images of a person taken from orthogonal views [2]. They adopted a generic 3D human model to represent both the human shape and articulation structure. Mapping 2D-silhouette information from the orthogonal view colour images onto the generic 3D model captures the shape of a specific person. Colour texture mapping is achieved by projecting the set of images onto the deformed 3D model. Fua *et al.* and D'Apuzzo *et al.* fitted animation models to the surface data derived from multi-image video sequences and extracted motion sequences for modelling [3, 4]. Lee *et al.* individualised a 3D facial model from two orthogonal pictures taken from front and side views [5]. This approach is based on the semi-automatic extraction of features from a face and modifying a generic facial model with the detected feature points.

All of the above methods achieved photo-realistic model appearance or believable human body animation motion but none of these approaches produced a highly realistic body shape for a specific individual. In the approach we report here we have made use of accurate facial and body 3D data (using C3D-based systems better than 0.5mm RMSE (Root Mean Square Error) can be achieved for 3D face capture and better than 2.0mm RMSE for 3D body data) to conform generic face and body models with what would appear to be a substantially greater degree of subjective fidelity (and realism) than the methods reported above.

In section 2, we describe our approach to segmenting the 3D human body surface into principal body components. Segmentation is an essential prerequisite to establishing global correspondences between the generic model and the captured data. Following in section 3; we describe a surface conformation procedure to transform the captured 3D data into an animatable form. Finally, we present results generated by our process for individualising a generic animation model based on real-world 3D human body data.

2. Segmentation

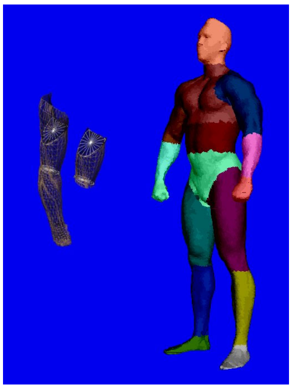

Figure 1: Landmark-based segmentation.

Usually 3D human body data captured by 3D imaging techniques contains little semantic information, i.e. the data comprises list or mesh of 3D data points without any indication of what body component they represent. At early stage of our research work, a technique for landmark-based, interactive human body segmentation had been developed such that the major body parts, comprising limbs, torso and head could be located and labelled. This system [6] enabled a user to specify limb and head segmentation planes by means of a user interface through which landmarks could be placed interactively on both the 3D imaged models and generic models (Figure 1)

A fully automatic body segmentation procedure has also been implemented [7] but has not yet been integrated into the conformation process described here. In this scheme the unstructured human body data is first sliced into contours from the top of the head to the soles of the feet and the limbs and torso are segmented based on topology of the contours (Figure 2).

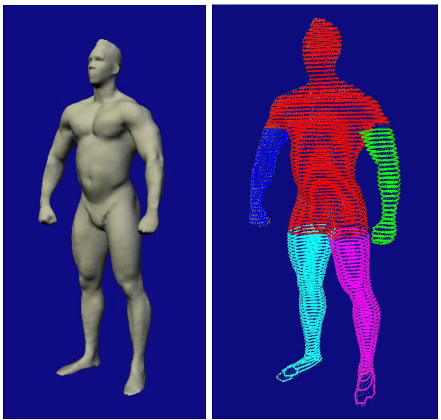

Figure 2: A CyberWare® body model (a) and automatically segmented result (b).

Once the generic model has been adjusted into the same posture as the captured data, the segmentation procedure is also applied to the generic model in order to break it into the same principal body parts (Figure 2) that correspond to those of the segmented 3D imaged body data. Global correspondences between the generic model and the captured data can then be established based on the groups of contours. These global correspondences enable the generic model to be conformed to the input 3D imaged body data.

3. Conformation

The conformation process comprises a two-step process: global mapping and local deformation. The global mapping registers and deforms the generic model to the scanned data based on global correspondences, in this case manually defined landmarks. The local deformation reshapes the globally deformed generic model to fit the scanned data by identifying corresponding closest surface points between them and then warps the generic model surface in the direction of the closest surface.

3.1. Global mapping

Global registration and deformation are achieved by means of a 3D mapping based on corresponding points on the generic model and the scanned data. The 3D mapping transforms the manually defined landmark points on the generic model to the exact locations of their counterparts (also defined manually) on the scanned data. All other points on the generic model are interpolated by the 3D mapping; this mapping is established using corresponding feature points as follows:

Let us denote two sets of corresponding 3D points given by $\mathbf{t}_i^T = (t_{i1}, t_{i2}, t_{i3})$ and $\mathbf{u}_i^T = (u_{i1}, u_{i2}, u_{i3})$, $i = 1 \dots n$ as the original and new feature points, respectively. We wish to deform our mesh so that point $\boldsymbol{t_i}$ moves to a new position $\boldsymbol{u_i}$ for all n feature points and then to interpolate the reminder of the 3D data set subject to these constraints.

A three dimensional deformation can be achieved by using three independent interpolations from $IR^3 \rightarrow IR$ to model the displacement of $\boldsymbol{t_i}$ (t_{i1}, t_{i2}, t_{i3}) along directions 1, 2, 3, (corresponding to the orthogonal x,y,z axes respectively) independently. If we denote these solutions $f_1(t), f_2(t), f_3(t)$ respectively, using the constraints:

$$f_1(t_{i1}) = u_{i1},\ f_2(t_{i2}) = u_{i2},\ f_3(t_{i3}) = u_{i3}$$

we can now define a deformation of any point in the mesh by:

$$\mathbf{t}^T \rightarrow (u_1, u_2, u_3)^T = (f_1(t), f_2(t), f_3(t))^T$$

This will be an exact interpolation for the points $\boldsymbol{t_i}$ and a smooth deformation elsewhere.

Using a radial basis function, σ, the solution of the 3D mapping takes the following form:

$$f_l(t) = a_1(l) + a_2(l)t_1 + a_3(l)t_2 + a_4(l)t_3 + \sum_{j=1}^{n} b_{jl}\sigma(t - t_j), \quad l = 1,2,3$$

with coefficients given by the solution of a system of linear equations. The radial basis function used is: $\sigma(t_i - t_j) = |t_i - t_j|$.

The global mapping results in the mesh of each body component of the generic model being subject to a rigid body transformation and then a global deformation to become approximately aligned with the scanned data.

3.2. Local deformation

Following global mapping, the generic model is further deformed locally, based on the closest points between the surfaces of the generic model and the scanned data. The polygons of the generic model are displaced towards their closest positions on the surface of the scanned data. Since the body models have been segmented, it is possible to avoid making erroneous closest point associations between points within the surfaces of the limbs and the torso. In order to speed the process of identifying the closest points, the polygons of the scanned data are organised within a Binary Space Partition (BSP) tree.

4. Results

Currently, the two-step model conformation process has been validated on a small number of body 3D images and head 3D images collected by the Cyberware laser scanning process, the C3D photgrammetric 3D imaging technique and also using the Wicks & Wilson Moiré-based approach to 3D imaging. A complete example of the process using Cyberware data is illustrated in figures 3-5. Figure 3a shows the original generic model and figure 3b shows the result of the initial 3D mapping process.

Having established global correspondences between the generic model and the 3D captured data, through the landmark-guided segmentation procedure, the generic model has been deformed globally to align with the 3D captured data. Exact correspondence between landmarks has been maintained while the positions of all other points have been interpolated. Comparing the 3D imaged model data (figure 4b) to that of the generic model (figure 3a), the global deformation of the initial 3D mapping has now brought the surface of the generic model into approximate alignment with the surface of the 3D data (figure 3b).

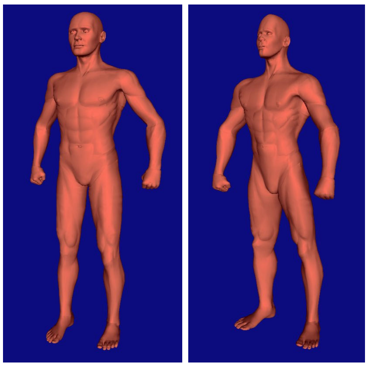

Figure 3: Original Generic model (a) and Generic model after the initial global 3D mapping (b).

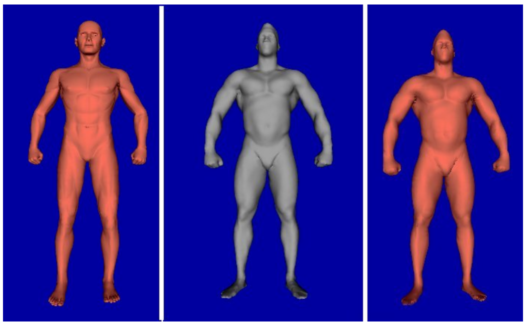

Figure 4: Generic model (a), 3D imaged body data (b) and final conformed result (c).

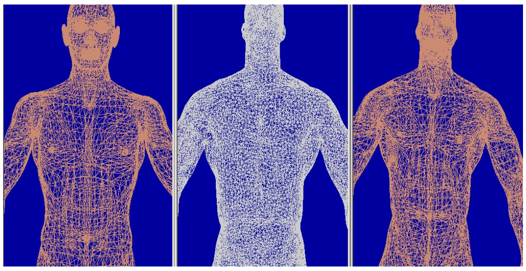

Figure 5: Generic mesh (a), 3D imaged body mesh (b) and final conformed mesh (c).

The polygons of the globally deformed generic mesh are further deformed to their closest positions on the surface of the captured data. Figure 4c shows this final conformation result, rendered using smooth shading.

Figure 5 shows the magnified wire-frame body surface representations to illustrate the differences in mesh topology between the captured human body data mesh, figure 5b, and the conformed generic model mesh, figure 5c. The conformed generic model has the same mesh topology as the generic model, figure 5a, but has the individualised shape, i.e. topography, of the 3D imaged real-world data. An animation sequence comprising a conformed individual taking a “virtual walk” is presented in figure 6.

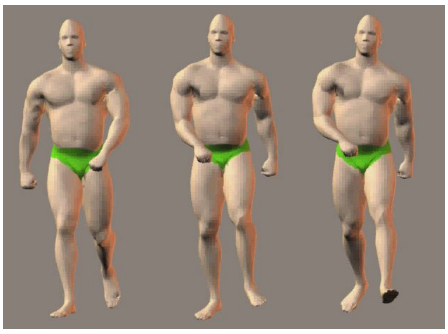

Figure 6: Animated individual.

5. Conclusions

The principal contribution of this work is to facilitate the creation of realistic 3D animation models of specific individuals by exploiting accurate 3D geometry of the individual captured by means of 3D imaging techniques. Through the adoption of a two-step conformation process we have first approximately aligned a generic animation mesh to the 3D imaged data by means of landmark-based limb segmentation and a global 3D mapping. Through this initial registration process we have brought the generic and 3D imaged data into sufficiently close alignment to allow a simple closest-point deformation algorithm to achieve close alignment of the two surface meshes. The input 3D imaged data with little semantic information has in effect been “cloned” by the conformed generic model. The resulting virtual human model comprises an inherently animatable mesh topology modulated with a highly realistic individualised body shape.

Acknowledgement

Thanks are due to Cyberware® of Monterey, California, USA for their excellent 3D whole body scanned data and, Engineering and Physical Sciences Research Council/Department of Trade and Industry for supporting the project through the *Imaging Faraday Partnership*.

References

1. J. P. Siebert and S. J. Marshall, “Human body 3D imaging by speckle texture projection photogrammetry”, *Sensor Review* **20** (3), pp218-226, 2000.
2. A. Hilton and T. Gentils, “Capturing human models to populate virtual worlds”, *CVSSP Technical Report*, June 1998.
3. P. Fua, A. Gruen, R. Plänkers, N. D'apuzzo, D. Thalmann, “Human Body Modelling and Motion Analysis From Video Sequences”, *Proc. International Symposium on Real Time Imaging and Dynamic Analysis*, Hakodate, Japan, June, 1998.
4. N. D'Apuzzo, R. Plänkers, A. Gruen, F. Fua and D. Thalmann, “Modelling Human Bodies from Video Sequences”, *Proc. Electronic Imaging*, San Jose, California, January, 1999.
5. W. S. Lee, N. M. Thalmann, “Fast head modelling for animation”, *Journal Image and Vision Computing*, **18**(4), pp355-364, 2000.
6. P. Siebert and X. Ju, Realistic human animation using scanned data, Proc. 5th Numérisation 3D/Scanning 2000 Congress, Paris, France, 2000.
7. X. Ju, N. Werghi and J. P. Siebert, “Automatic Segmentation of 3D Human Body Scans”, *Proc. IASTED Int. Conf. on Computer Graphics and Imaging 2000* (**CGIM 2000**), Las Vegas, USA, 2000.