

# Body scanning for avatar production and animation

Andrew Brownridge and Peter Twigg

*Faculty of Science & Engineering, Manchester Metropolitan University, Manchester, UK*

*Work presented in this report describes a workflow developed to produce and animate a realistic virtual human for animation. Production of the avatar makes use of infrared depth sensing technology to produce a representative avatar of a scanned subject. The potential is that this could be used to contribute to a virtual fashion show or for motion analysis purposes. Several methods of constructing a clean polygonal character mesh from the scan data are trialled and discussed. The preparation of the resulting character mesh for animation involves rigging techniques used within three-dimensional (3D) animation and are optimised using scripts developed for this specific task. The report investigates how extracted measurement data can be used to place joints as a starting point for the rigging process. Pre-recorded motion capture data are successfully used to animate a representative avatar of a human subject in a realistic and natural way within a virtual 3D environment.*

*Keywords: motion capture; body scanning; avatar*

This paper presents a workflow for the production of a virtual character representing a human subject, which will be used for animation through motion capture. The work contributes to a project, which aims to record the motion of ballet dancers for both performance analysis and corrective coaching, or augmented presentation of dance. In both cases, it would be required that an avatar representation of the subject be displayed in a real-time virtual environment. A body scanner commonly applied to clothing design and sizing standards development in fashion is used for the production of a polygonal actor mesh. Autodesk's Maya, a three-dimensional (3D) modelling and animation software, is used for character rigging. Data-processing methods are investigated to establish appropriate techniques for producing a clean character mesh and a joint hierarchy suitable for animation methods. The results show that a fully rigged character can be created efficiently for the rapid production of a virtual dancer for animation in real-time using motion capture technology. The same techniques can be applied within other disciplines to offer augmented performances such as a virtual fashion show, which has been previously attempted (Okada, Stenger, Ike, & Kondoh, 2006).

## Previous work

A wireless, inertial gyroscopic sensor-based, system was used to record the movements of ballet dancers to investigate if such an equipment is capable of detecting misalignments in ankle position (Hiller, Refshauge, & Beard, 2004) and for establishing real-time recording and virtual performance (Andreadis et al., 2010; Meador, Rogers, O'Neal, Kurt, & Cunningham, 2004). Motion capture techniques allow motion, usually human, to be recorded, for presentation and analysis, and associated technologies are becoming increasingly sophisticated, accurate and portable (Menache, 2000). Two broad approaches are commonly used: electromagnetic and optical tracking, which each have their benefits and constraints (O'Rourke, 2003; Parent, 2008). More recently, however, scanner-based systems have been investigated (Berger et al., 2011; Shotton et al., 2011).

Body scanning technology is becoming more widely used throughout the fashion industry in garment design (Gill & Hayes, 2012) and body shape analysis (Grogan, Gill, Brownbridge, Kilgariff, & Whalley, 2013). Previous research has been carried out to explore the implementations of combining body scanning technology and motion capture systems within the apparel industry (Zong & Lee, 2011).

## Data recording

Work described in this paper is based on data recorded with the Animazoo IGS-190 inertial motion capture suit. The suit makes use of 18 small solid-state inertial sensors (gyros) to measure the rotations of the actor's joints in real-time. The requirements of the subject were that they could perform in their usual environment in an unhindered way, and the requirements of the equipment were to provide a representation of the subject in real-time and to record sufficient and accurate movement data for analysis. A Biovision Hierarchy (BVH) data file – a format originally developed by Biovision – provides skeleton hierarchy information in addition to the motion data. This is the industry standard file format for motion capture data. The location for the initial dance testing was the Northern Ballet School. The motion data are able to be viewed in real-time as a simplistic skeletal representation in software developed by Animazoo, or as an animated avatar through software developed by Autodesk (see Figures 1 and 2). The motion data can be recorded and output to BVH files for later analysis. The dancers reported that the IGS-190 suit had no restrictive effect on their movements and after a 5-minute familiarisation-warm-up routine, they were not aware of the presence of the equipment. The first session showed that the suit is capable of recording the extreme motion associated with ballet and that the equipment could be accurately set up and calibrated in a relatively short time, typically less than 15 minutes. In order to produce a realistic real-time operation avatar, a suitable accurate character mesh is required with accurate skeleton rigging and realistic UV skin mapping.



Figure 1: Dancer practising with the motion capture suit

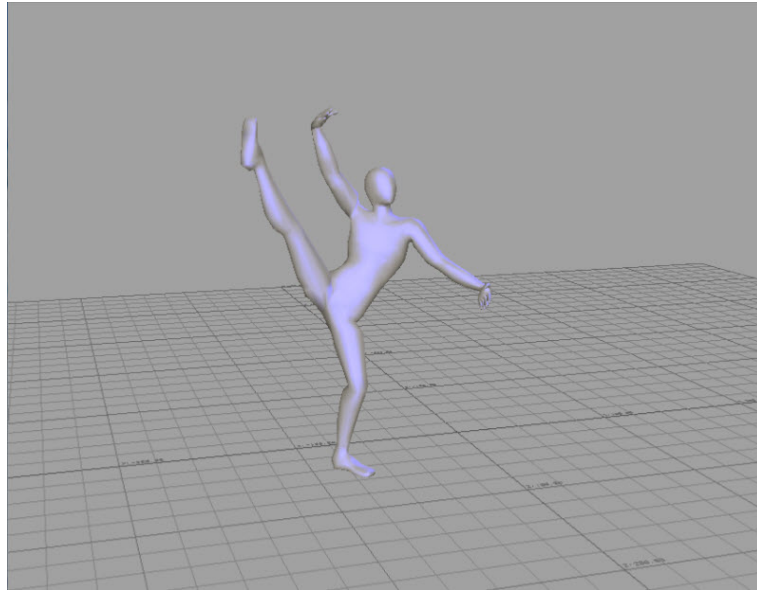


Figure 2: Real-time data visualisation of the dancer in Autodesk's MotionBuilder

## Body scanning

To produce a mesh that accurately represented the actor, a body scanner was used: namely a [TC]2 KX-16. The body scanner offers technology which is commonly applied within the fashion industry for collecting measurements of the physique (Li & Li, 2010), which is then analysed for the purpose of garment design (Tyler, Mitchell, & Gill, 2012; Xu, Huang, Yu, & Chen, 2002) and sizing standards development (Loker, Ashdown, & Shoenfelder, 2005; Wren & Gill, 2010). The non-contact, 3D measurement system is based on infrared depth sensing technology similar to that used in the Microsoft Kinect (Doi, Shimizu, & Sato, 2007; Tong, Zhou, Liu, Pan, & Yan, 2012) and uses distance information from the depth sensor to produce a digital copy of surface geometry of the scanned object, similar to methods used for model reconstruction in archaeology (Doi et al., 2007). 3D images are taken of the body within seconds, generating a silhouette of the shape and an extensive list of actual body measurements (see Figure 3).

The test scanning session was planned for a midweek lunchtime and it was anticipated that it would take no longer than one hour to acquire the necessary data. Upon arrival at the scan site, it became apparent how compact the setup was. The scanner was no larger than a changing cubicle and a desktop PC was connected to the side. The scanning process for taking measurements of the body requires the subject undresses down to underwear and that the underwear was of a neutral colour (grey, beige, etc.). The subject to be scanned proceeded to change ready for the scan whilst the PC was booted. Three scans were taken in underwear, each taking eight seconds to complete. A further scan was taken in full clothing to examine whether the scan would produce a usable mesh for the animation of a clothed virtual character. Once the scan had completed after eight seconds, the data were immediately available for viewing (see Figure 4). The scan data were available in a number of different file formats: .rbd (point cloud data), .wrl (VRML text data file) and .bin (depth scan image). The software also had a function that would use the scan data combined with a template polygon model of a human body to produce a smooth mesh based on the measurements taken from the scan. This was the output in .obj file format.

The whole process took less than 30 minutes from arriving at the site to having the output files on a USB drive ready to take away. This process would be quicker with greater efficiency and experience of the workflow, and a large number of subjects could be scanned in a short period. The scanner itself can be made mobile with a set-up and collapse time of only a few hours.

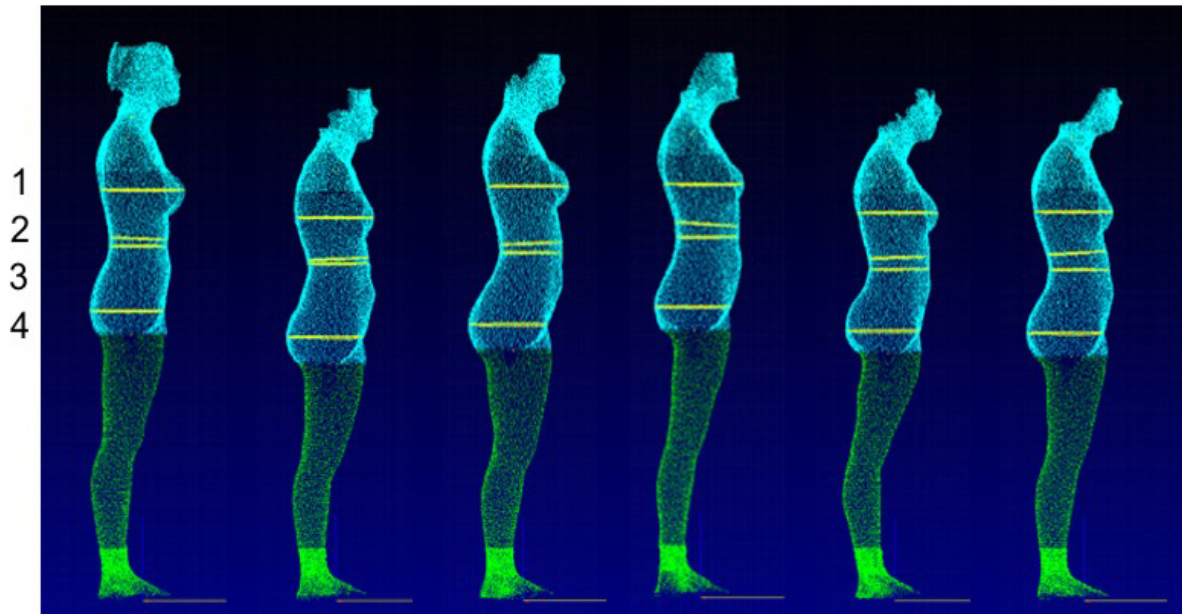


Figure 3: Scan data used for garment design (Tyler et al., 2012)  
Six subjects who would all be classified as size 12 from recent industry size specifications  
i.e. Measurements fall within range: Bust 89-93 cm, Waist 71-75 cm, Hip 96-100 cm.  
Subjects are between 19-22 years.  
Measurements shown are: 1. Bust, 2. Waist A (natural indent), 3. Waist B (small of back), 4. Hip

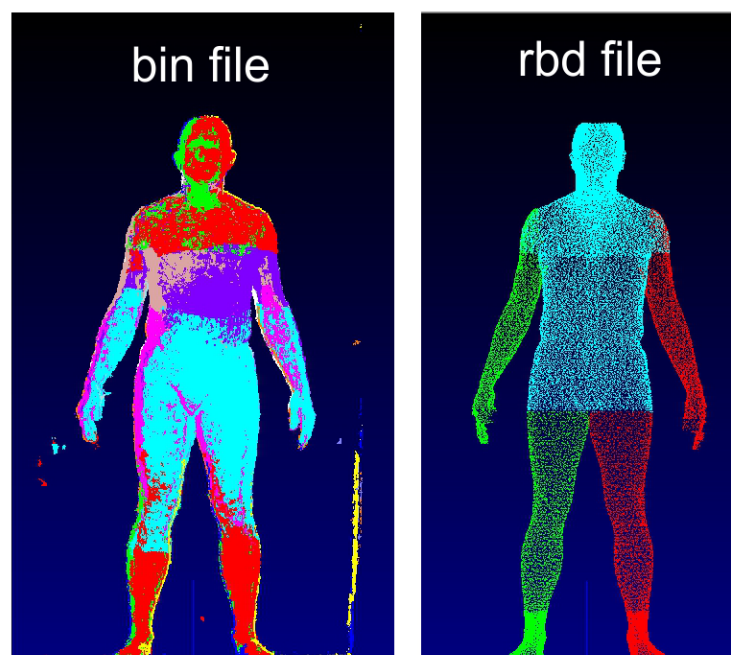


Figure 4: Depth scan image and point cloud data

## Preparing mesh data

Once the scan data had been acquired, assessment was required to establish the most suitable method of producing a polygonal mesh that accurately represented the subject and could be animated by means of motion capture data with realistic deformation. A number of techniques were explored within a range of software using the different output files from the scan session.

The first attempt at reproducing a representative mesh of the subject used a piece of software called MeshLab. The software had the ability to import the .wrl and .obj file formats from the scan data acquisition (see Figures 5 and 6). When importing the .wrl file format, the software takes the point cloud data points contained within the .wrl file (in coordinates in the x-, y-, and z-axes) and joins the points together forming edges and polygonal faces. There are a number of tools within the software which can be used to attempt to 'clean up' the mesh such as: removing duplicate or null faces, merging close vertices, and flipping faces whose normals are opposite to all of its adjacent faces. The software was found to be somewhat unreliable and had the tendency to crash when multiple operations were performed. It did, however, have the ability to export the mesh to different file formats such as .ply, .obj, and .u3d, which could be recognised by other software.

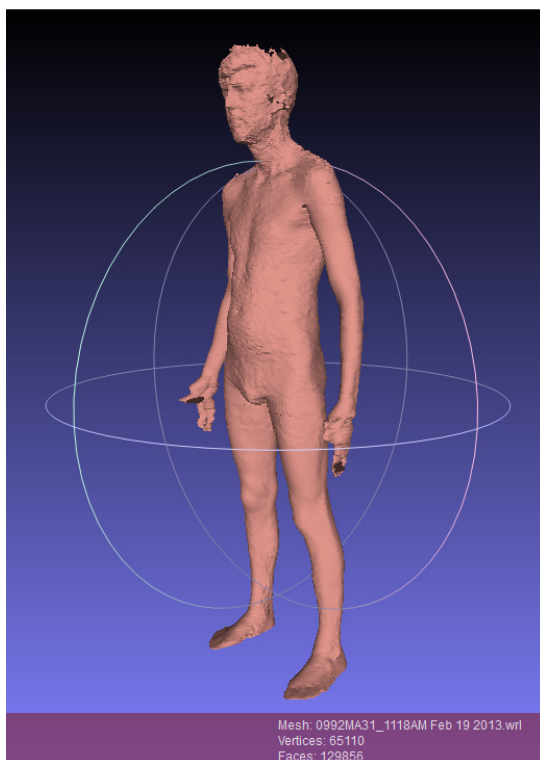


Figure 5: MeshLab wrl import

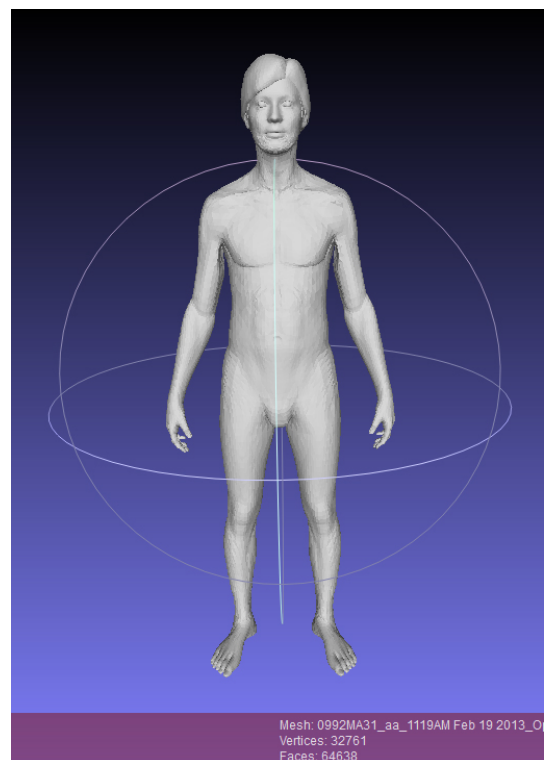


Figure 6: MeshLab obj import

The second piece of software that was used was Pixologic's ZBrush; this had the ability to import the .obj file from the data collected at the scan session, and the .wrl file converted to the .obj file through MeshLab (see Figures 7 and 8). ZBrush is a powerful sculpting tool used extensively for character creation. It was anticipated that the rather rough looking .wrl conversion could be smoothed using the tools within ZBrush, and that the inbuilt retopologising plug-in 'QRemesher' would allow either mesh to be optimised for animation.



The smoothing of the converted .wrl data was a painstaking process, which required a great deal of skill and time, and was subject to deviation from the measurements taken from the scan of the subject. The results were not as visually pleasing as the smoothed .obj file produced from the scanner software and, dependant on the artist that sculpts the converted .wrl file, is not as accurate a representation of the subject. In addition this is a much longer process; in contrast the smoothed .obj file is available almost instantly from the scanner software.

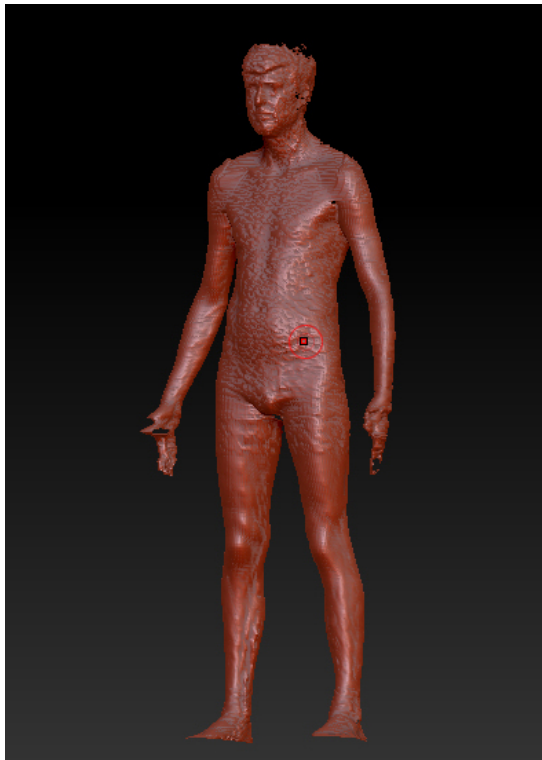


Figure 7: ZBrush wrl-obj import



Figure 8: ZBrush obj import

It was concluded from this that the most suitable workflow for creating a representative polygonal mesh of a subject was to use the software supplied with the scanner. Although it lacks the subtle individual characteristics of the subject, it quickly produces a proportionally accurate representation. It would take a very skilled 3D artist a long time to produce a model that would contain the subtle individual characteristics of a subject and potentially at a cost of a less proportionally accurate model. This is not to say that this method does not have its issues. As it was evident from the .wrl import to MeshLab, there are problems accurately scanning the hands, feet and head. This can be explained because of the positioning of the sensors and their inability to capture the intricate details between the fingers, and the limitations of scanning the base of the feet and top of the head. The software that creates the smoothed mesh from the scan data tackles this problem by using pre-made hands, feet and head models and attaches these to the body mesh. This results in the head, hands and feet not being actual representations of the subject's relative surface geometry and there is a split in the combined mesh at the neck. The software also attaches a hairstyle to the mesh. The separated elements of the head and hairstyle from the neck split can be seen in Figure 9.

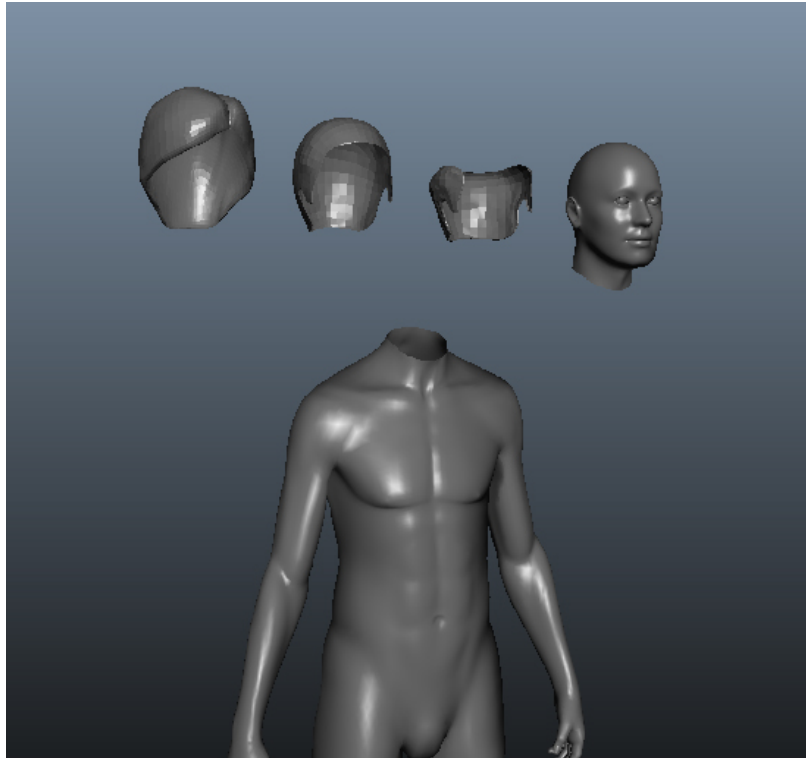


Figure 9: Separated parts of the combined mesh

## Preparing the character mesh for animation

In order to correct the issues described above, the mesh can be imported in the .obj file format into Autodesk's Maya software. There were minor surprising results in the import of the mesh to Maya. It appears that the global axis and scaling units differ in the scanner software and Maya, as the character was imported sideways into Maya and was relatively very large. The character had to be resized and oriented to fit the Maya scene with rotations of  $-90^\circ$  in the x- and y-axes and scaling at 0.01 in the x, y and z (effectively scaling from metres to centimetres). As this would be an inconvenient, yet necessary step with the import of any character from a scanned obj mesh, a work around was developed in the form of a Maya-embedded language (MEL) script.

The neck split can be stitched and the edges softened using a further MEL script which merges all vertices at a specified distance apart and makes soft an edge if the angle between two owning facets is flatter (smaller) than the specified smoothing angle. Any further problematic geometry can then be corrected manually or by using a clean-up tool within the software.

Along with the smoothed polygonal mesh produced by the software, a UV map is also produced for the model (see Figure 10). This is particularly useful for texturing the model. It can be quite an arduous task laying out a model's UV map, particularly on an organic model so complex. To have this UV map produced to compliment the model saves a lot of time and would enable a quick turnaround mapping a photograph of the subject to the virtual characters' mesh.

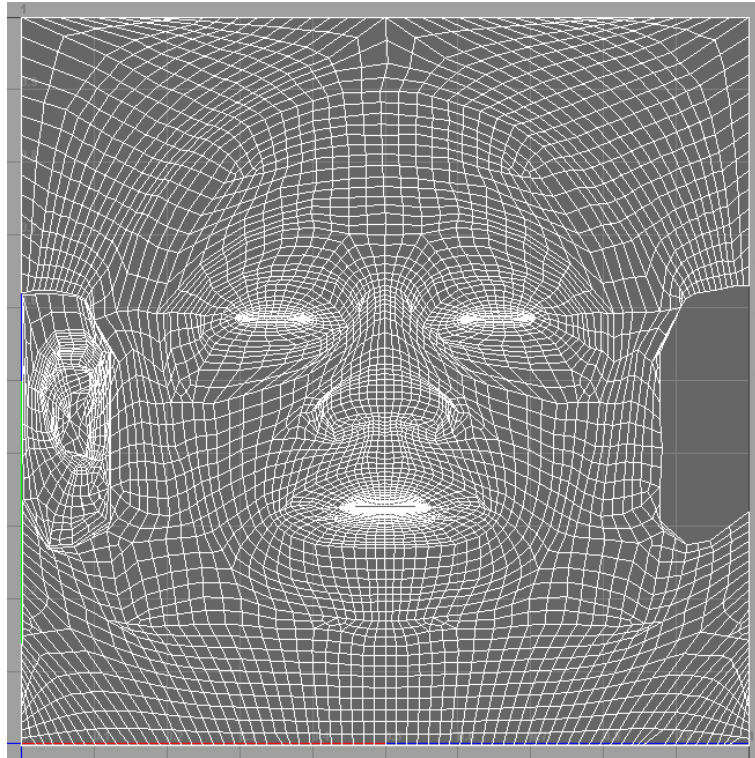


Figure 10: UV map of face

## Rigging for animation

Once the character shell was complete, the next stage was to develop a rig with a joint hierarchy suitable for animation by motion capture. The conventional way of doing this is with the character in a 'T-pose', which describes the pose of a character whose arms are stretched out to the side lying horizontally. This pose makes the character design, rigging and animation processes more straightforward in 3D space. However, because of the way the body scanner worked it was necessary to do this with the character shell in a pose with the arms by its side. This resulted in a number of problems. For instance, when placing the joints for the arms and hands with the character in a T-pose, it is a simple process to view from the top orthographic view and work down the arms automatically snapping to the base of the grid (y-axis at 0). Once the joint chain has been aligned from the top view, it can simply be translated up the y-axis to fit into the geometry. With the character's arms relaxed to the side of the body, this process was not possible. Each joint had to be placed within the geometry using a combination of the views resulting in a time-consuming and intricate process, particularly for the fingers. The finished skeletal hierarchy can be seen in Figure 11.

No motion data are recorded for the fingers in the current workflow; however, joints were added to the fingers in order to enable the possibility of adding secondary animation to the hands such as the flexing and curl of the fingers. This is not necessary for the investigation of joint alignment in the lower limbs of ballet dancers, but the workflow was explored to determine the suitability of other sports or performing arts applications for future research.

As part of the scanning process, the software that accompanies the scanner records a number of measurements of the subject. From these measurements, the number of joint locations is calculated



and stored within a text file in the output folder of the scan. The first line of the text file reads 'l\_hip - 15.22 127.56 952.51

---

. Separated by a space, the first part is the name of the joint followed by the co-ordinates in units of centimetres in the order: front, left and up.

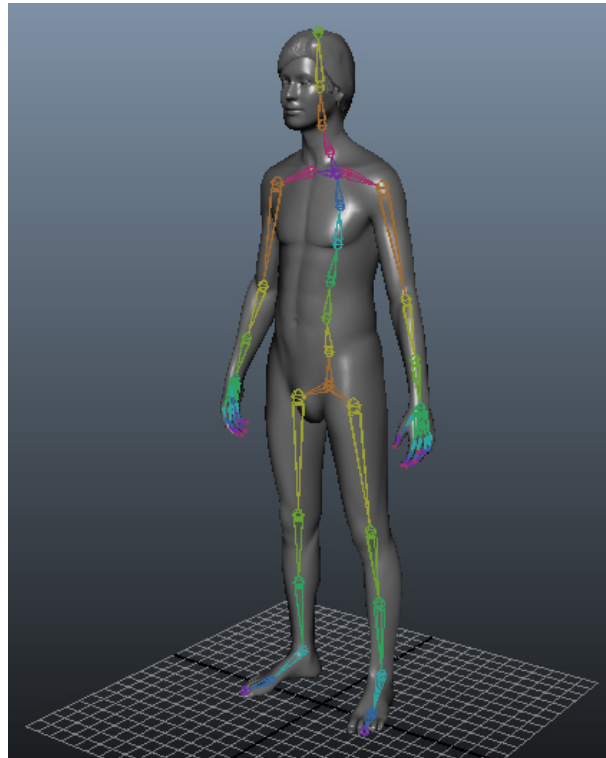


Figure 11: Skeletal rig

Maya has the ability through Python scripting (high-level programming language) to read from a text file and use the data to place objects within the scene. It was considered that it might be possible to use this joints file to place the joints for the character rig. A script was written to place joints at the locations specified in the text file. It can be noted that the orientation axis and scaling differences between the software have been addressed within the script by reordering the zxy from the scanner (front, left, up) to the xyz order used in Maya and converting centimetres to metres by division of 100. A significant problem was encountered from the results of the initial script, in that the joints were parented in the order that they were introduced (as is the case when constructing a joint chain in Maya). The script was subsequently amended to place the joints with no hierarchical chain by clearing the selection between creations (since creating a new joint leaves it selected, subsequent creations are parented to the previous one). The results of the former and latter mentioned scripts can be seen in Figures 12 and 13, respectively.

Although the scripts do not construct the ideal skeleton hierarchy for the purpose of animation through motion capture, they provide a good starting point for the placement of joints saving much time in the rigging process.

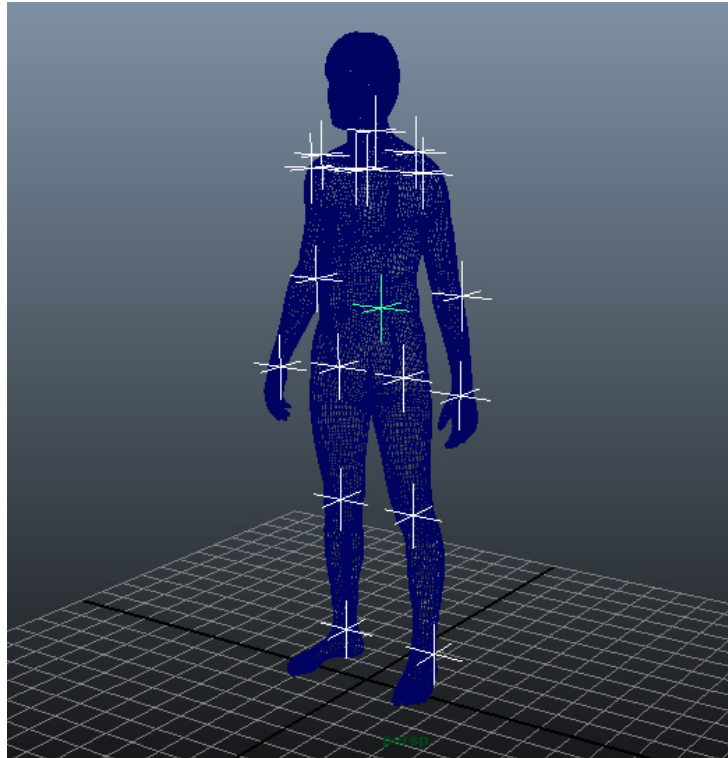


Figure 12: Locators places at joint positions

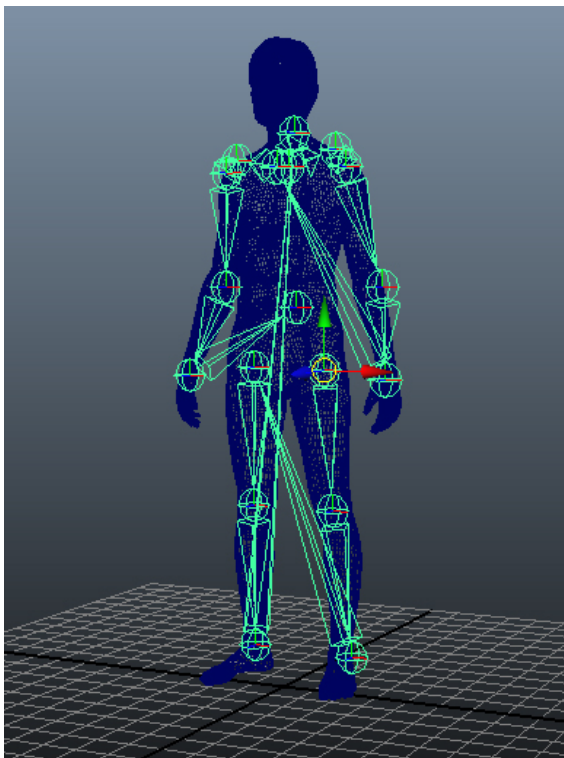


Figure 13: Chain of joints drawn

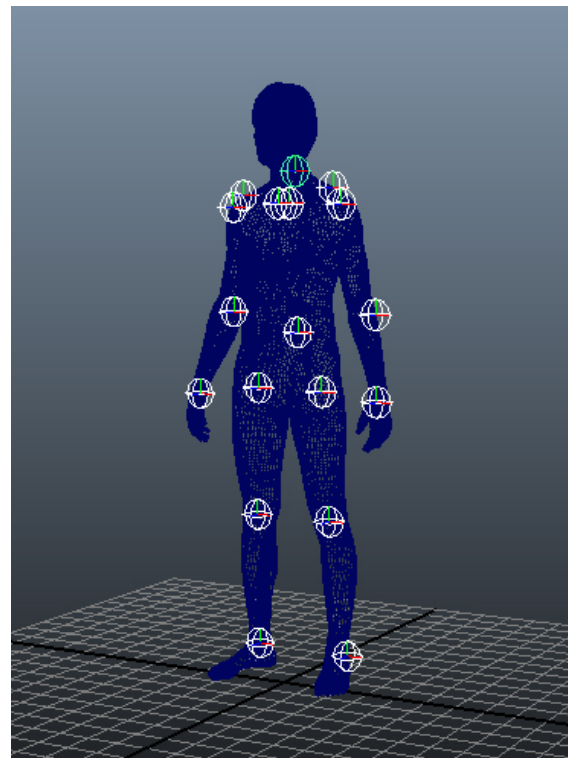


Figure 14: Individual joints placed

Once the skeleton hierarchy had been constructed, it needed to be bound to the geometry in order for the geometry to move and deform through retargeted animation. Usually, adjusting a deformable

mesh to achieve plausible results is a relatively straightforward process. Again, due to the character having its arms by the side rather than in the conventional T-pose, this proved somewhat problematic. When the shoulder joint was rotated to move the arms, it became evident that the close proximity of the arm and hand joints to the body had caused an association with the torso of the character. As the arms moved up, the sides of the character followed, which can be seen in Figure 14. This had to be manually corrected using a combination of tools within Maya and was once again a time-consuming and intricate process. The mesh deformation profile for the character can be saved to an XML data file with the weighting values corresponding to the vertex index. As the avatar is generated from a template mesh, subsequent avatars will share the same topology and the weighting values can be mapped onto a new character using this XML file. The mesh with corrected deformations can be seen in Figure 15.

The rigged character was exported to an empty scene within MotionBuilder, another Autodesk product within the entertainment creation suite, along with a pre-recorded motion capture file in the BVH file format (a file providing skeletal hierarchy information as well as the motion data). MotionBuilder reads the skeletal definition information contained within the header section of the BVH file and builds a representative model. In order for the BVH file to drive the animation of the avatar, the motion data need to be retargeted to the input of the avatar. The results of the motion data driven avatar can be seen in Figure 16.

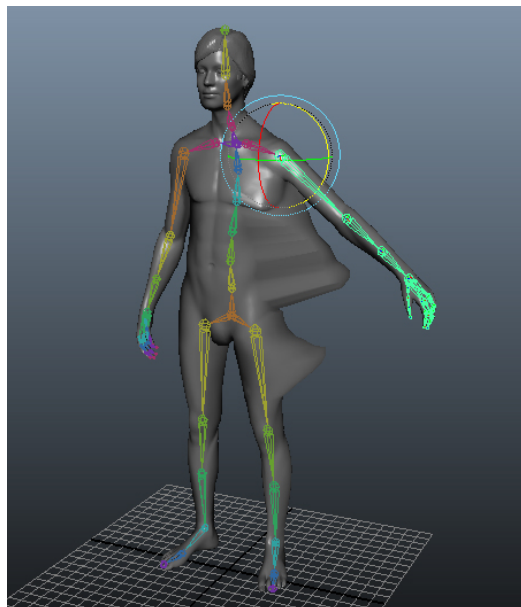


Figure 15: Default skin weighting

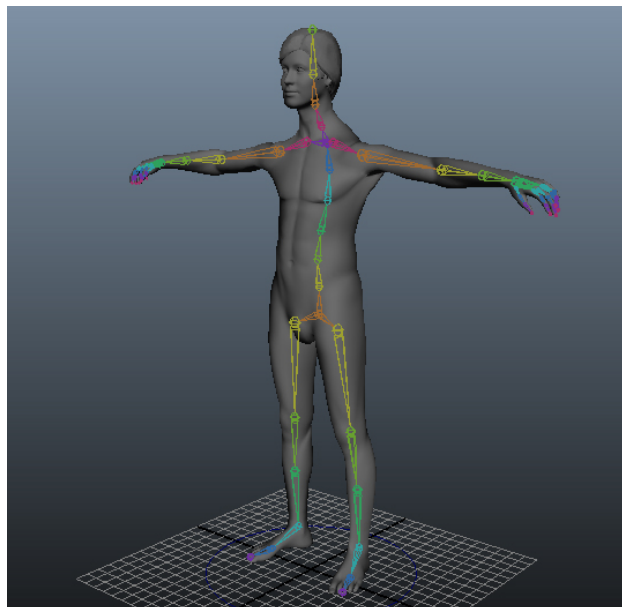


Figure 16: Adjusted skin weighting

## Conclusion and further work

The above investigation describes the workflow used to create a virtual character, which accurately represents the form of a scanned subject. This avatar can then be animated using motion recorded from the same subject using the IGS-190 motion capture suit and viewed in real-time within a virtual environment.

It has been shown that a proportionally accurate character avatar with clean topology suitable for animation can be created using depth sensor body scanning. A template character mesh was sized according to measurements of a subject calculated by the [TC]2 KX-16 software. This mesh was then cleaned up for animation using MEL scripting. A joint hierarchy was placed within the mesh combining a Python script which identified and placed joints at locations taken from the body scanner and manual rigging methods. The character mesh was then bound to the skeleton and geometry deformations were controlled by manual skin weighting methods. The skin weighting profile was saved to a data file for use in future avatar development. The rigged character was then animated within a virtual environment using inertial motion capture of a human performer.

Work will focus next on producing realistic and accurate avatar representation for real-time performance in a virtual environment and the development of performance analysis tools for corrective coaching. This can then be applied within a number of disciplines including fashion design, the performing arts and sports.

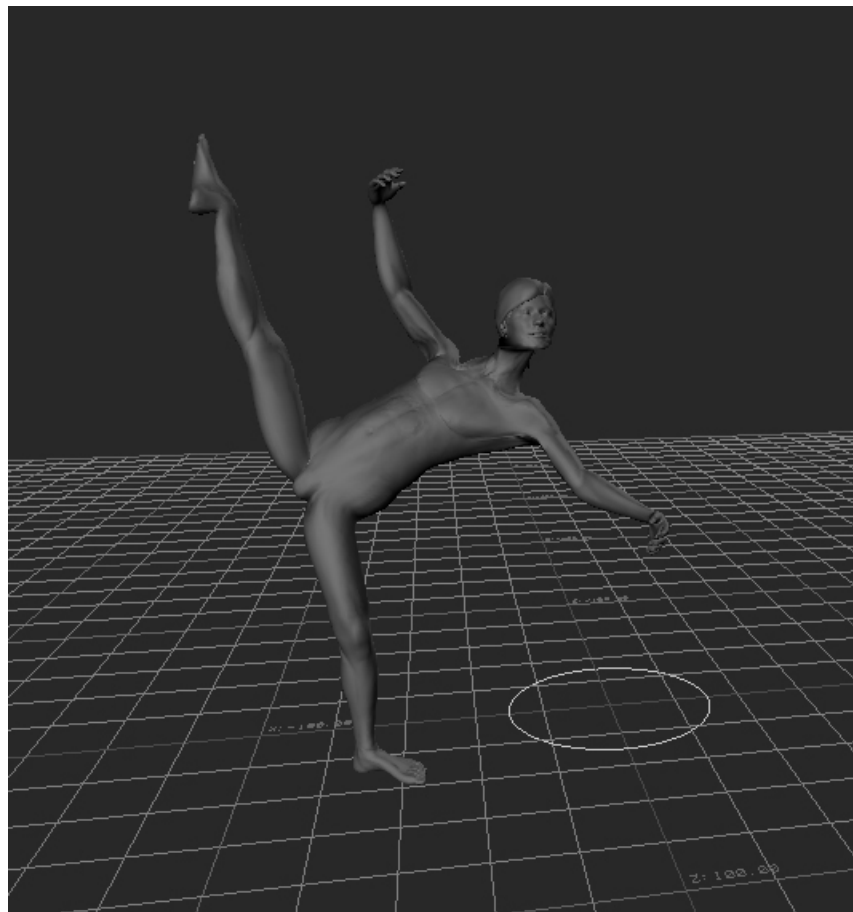


Figure 17: Avatar driven by dance capture

## References

- ANDREADIS, A., HEMERY, A., ANTONAKAKIS, A., GOURDOGLOU, G., MAURIDIS, P., CHRISTOPOULOS, D. & KARIGIANNIS, J. N. (2010, September 10-12). *Real-Time Motion Capture Technology on a Live Theatrical Performance with Computer Generated Scenery*. 14th Panhellenic Conference on Informatics (pp. 148 – 152). Tripoli, Greece.
- BERGER, K., RUHL, K., BRÜMMER, C., SCHRÖDER, Y., SCHOLZ, A. & MAGNOR, M. (2011, October 4-6). *Markerless motion capture using multiple color-depth sensors*. 16th International Workshop on Vision, Modeling and Visualization, Berlin, Germany.
- DOI, J., SHIMIZU, K. & SATO, W. (2007, May 7-9). *High-Density Scanning for Virtual Heritage and Archeology: Reproduction and Restoration*. 3DTV Conference (pp.1-4). Kos, Greece.
- GILL, S., & HAYES, S. (2012). Lower body functional ease requirements in the garment pattern. *International Journal of Fashion Design, Technology and Education*, 5 (1), 13–23.
- GROGAN, S., GILL, S., BROWNBRIDGE, K., KILGARIFF, S., & WHALLEY, A. (2013). Dress fit and body image: A thematic analysis of women's accounts during and after trying on dresses. *Body Image*, 10 (3), 380–388.
- HILLER, C. E., REFSHAUGE, K. M. & BEARD, D. (2004). Sensory Motor Control Is Impaired in Dancers With Functional Ankle Instability. *The American Journal of Sports Medicine*, 32, 216 - 223.
- LI, X. & LI, X. (2010). *Human Body Dimensions Extraction From 3D Scan Data*. International Conference on Intelligent Computation Technology and Automation, Changsha, China.
- LOKER, S., ASHDOWN, S. & SHOENFELDER, K. (2005). Size-specific analysis of body scan data to improve apparel fit. *Journal of Textile and Apparel, Technology and Management*, 4 (3), 1 - 15.
- MEADOR, W. S., ROGERS, T. J., O'NEAL, K., KURT, E. & CUNNINGHAM, C. (2004). Mixing Dance Realities: Collaborative Development of Live-Motion Capture In a Performing Arts Environment. *Computers in Entertainment*, 2, 12.
- MENACHE, A. (2000). *Understanding Motion Capture for Computer Animation and Video Games*. San Francisco, CA: Morgan Kaufmann.
- OKADA, R., STENGER, B., IKE, T., & KONDOH, N. (2006). Virtual fashion show using real-time markerless motion capture. In *Asian conference on computer vision*. Berlin: Springer-Verlag.
- O'ROURKE, M. (2003). *Principles of three-dimensional computer animation: modelling, rendering, and animating with 3D computer graphics*. New York, NY: Norton.
- PARENT, R. (2008). *Computer Animation - Algorithms & Techniques*. Waltham, MA: Morgan Kaufmann.
- SHOTTON, J., FITZGIBBON, A., COOK, M., SHARP, T., FINOCCHIO, M., MOORE, R., KIPMAN, A. & BLAKE, A. (2011). *Real-time human pose recognition in parts from single depth images*. Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society. Colorado Springs, Colorado, USA.
- TONG, J., ZHOU, J., LIU, L., PAN, Z. & YAN, H. (2012). Scanning 3D Full Human Bodies Using Kinects. *IEEE Transactions on Visualization and Computer Graphics*, 18 (4), 643 - 650.
- TYLER, D., MITCHELL, A. & GILL, S. (2012). Recent advances in garment manufacturing technology: joining techniques, 3D body scanning and garment design. In R. SHISHOO (Ed.), *The global textile and clothing industry: Technological advances and future challenges*. Cambridge, UK: Woodhead.
- WREN, P. A., Gill, S. (2010, November 2-3). *Industry Fit Practices and Their Variation*. Textile Institute 100th World Conference, Manchester, UK.
- XU, B., HUANG, Y., YU, W. & CHEN, T. (2002). Body Scanning and Modeling for Custom Fit Garments. *Journal of Textile and Apparel, Technology and Management*, 2 (2), 1 - 11.
- Zong, Y., Lee, Y.-A. (2011). An exploratory study of integrative approach between 3D body scanning technology and motion capture systems in the apparel industry. *International Journal of Fashion Design, Technology and Education*, 4 (2), 91–101.