



Archived at the Flinders Academic Commons:
<http://dspace.flinders.edu.au/dspace/>

'This is the peer reviewed version of the following article:

Wyeld, T., Hobbs, D. Visualising human motion: A first principles approach using Vicon data in Maya (2016) Proceedings of the International Conference on Information Visualisation, 2016-August, art. no. 7557930, pp. 216-222. DOI: 10.1109/IV.2016.83

which has been published in final form at
DOI:

<http://dx.doi.org/10.1109/IV.2016.83>

Electronic ISSN: 2375-0138

"Copyright © 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works."

Visualising Human Motion: a First Principles Approach using Vicon data in Maya

Theodor Wyeld, David Hobbs,
Flinders University

{ theodor.wyeld@flinders.edu.au, david.hobbs@flinders.edu.au }

Abstract

This paper describes a first principles approach to understanding how 3D digital animation of human motion can be processed and produced from raw data, without the use of proprietary software. The paper describes how students collected motion data using a custom marker set, how this was used to create a point cloud, how errors were corrected, and finally how the skeleton was rigged, skinned and modelled in Maya.

Keywords: 3D animation, human motion, Maya, Vicon, first principles, marker set, point cloud, modelling, MEL.

Introduction

At undergraduate level it is critical that students of 3D animation, visual effects or games development understand the fundamental principles of human motion capture. As such, a method was needed to provide this understanding. Access to a Vicon-equipped motion capture laboratory provided the basis for this project. The laboratory's main use is for clinical motion analysis studies.

The 3D media students that this project was aimed at use Maya 3D, among other Autodesk products, in their productions. Unfortunately, the Vicon dataset cannot be imported directly into Maya. Although it can be exported to other software, such as *MotionBuilder* or *Blade*, they create ready-formed rigs. As such, using software other than Maya 3D does not provide an understanding of the fundamental underlying principles involved in the use of motion capture data for animation that this project seeks to demonstrate.

Without a critical understanding of the processes and concepts involved in the use of motion capture data for animation, students are unable to solve problems arising from poorly constructed rigs or incomplete datasets. Hence, the emphasis with this project was to provide a fundamental understanding of an otherwise complex process – a process for which there is little background literature available. This is mostly due to the relative infancy of motion capture compared to the film, animation, and games industries, more generally. Furthermore, organisations that provide motion capture services to the industry appear reluctant to provide details of their processes. This is understandable as they have invested heavily in their proprietary procedures and charge accordingly. Hence, there is no incentive for them to make the details of their procedures public.

In part, this project attempts to address the gap in public knowledge of motion capture procedures for the benefit of undergraduate students. It outlines a procedure for:

- preparing to capture human motion using the Vicon suit of tools;
- generation of a data point cloud from the Vicon dataset in Maya;
- extrapolation of the joint locations from surface marker data;
- error correction; and,
- rigging and skinning in Maya.

This paper concludes with a discussion on the importance of understanding how to overcome problems using motion capture data.

Background

Capturing the motion of people and animals was first photographically investigated by Eadward Muybridge around the turn of the twentieth century. His photographic and cinematographic studies of humans and animals in motion explained many phenomena previously unknown. For example, his famous chronophotographic study of a horse revealed that all four of its hoofs are off the ground when in full gallop (see Figure 1). His studies also photographically mapped many biomechanical actions not possible using other methods (Adam, 2014).

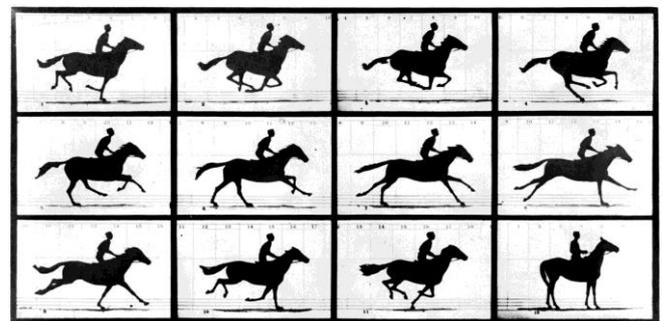


Figure 1. *The Horse in Motion* (Muybridge, 1878, Library of Congress Prints and Photographs Division).

Today, the capturing of human motion involves the integration of the fields of biomechanics, human movement, computer vision, signal processing and computer graphics. The interdisciplinary nature of this

research leverages the advances in each field, culminating in the study of motion capture more generally.

Motion capture can typically be divided into the two-dimensional and three-dimensional domains. Two-dimensional human motion tracking involves using a single or dual static camera setup and compares the silhouette of the human form against a background in a single plane. In studies of human locomotion this means capturing the frontal (coronal) and side (sagittal) views of human gait, which can be viewed individually or simultaneously using a split screen to study the gait pattern. This is similar to what Muybridge had achieved with his capture of a galloping horse for the sagittal view only. Three-dimensional points can be extracted from a single camera when the camera is moving in a scene. However, the results are not always reliable. Three-dimensional points are extracted using perspective projection into a planar displacement modelled as a two-dimensional transformation. Different arrangements of the various markers (used to capture the movement) relative to each other can generate the same two-dimensional projections. This can lead to ambiguity and inaccuracies in the resulting three-dimensional dataset which are difficult to resolve (Korc & Hlavac, 2008). Hence, three-dimensional motion tracking typically involves at least two spatially-calibrated cameras and often many more.

Motion capture is used in several industries, including animation, visual effects (VFX), game development, sports and rehabilitation, among others. The laboratory that this paper reports on was established to research upper and lower limb human movement. It utilises 10 Vicon (www.vicon.com) Bonita cameras suspended from scaffolding on the perimeter of an 8.5m x 10m room, with 4 AMTI (www.amti.biz) force plates submerged in the floor to measure ground reaction forces as subjects walk across them. The Vicon software is used to collect kinematic data from the cameras and kinetic data from the floor plates. It is installed on a standard Vicon issued desktop PC (Dell T5810, Intel® Xeon® CPU E5-1620 v3 @ 3.50 GHz, 16GB RAM, 64-bit OS, Windows 7 Professional). Subjects typically wear just their underwear during a clinical study, or tight clothing for a non-clinical study, with up to 40 reflective markers placed at strategic locations on the body according to a predefined anatomical marker set. Data is gathered in real-time and can be exported in various file formats, including ASCII, CSV, C3D, MKR, MP, TRC, VSK and XCP.

The purpose of this project was to extend the usability of the University's motion analysis laboratory to service motion capture other than for clinical or rehabilitation purposes alone. This included the need to generate datasets suitable for animation, VFX and game development. These typically require a different type of data for generating animated motions than those provided by the Vicon software. Where the Vicon software generates surface animated motions based on surface markers there was a need to generate pivot-point

animations, more suitable for animation, VFX and games development. Pivot-point animation relies on the rotation of pivot points aligned with the intersecting bone joints of the human skeletal system. This is inherently different to surface animation which relies on muscular contractions and distortions at the surface of the human body used in gait analysis. As such, a method was needed to extract pivot-point rotations from the surface data.

Gait Analysis

Within a clinical setting, health professionals have been observing, analysing and monitoring abnormal and impaired human locomotion for rehabilitation purposes for decades. Observational gait analysis – as the name suggests, using the naked eye to observe an individual's gait for diagnostic purposes – is the quickest, cheapest and usually first measure for identifying abnormal gait patterns compared to 'typical' or unimpaired gait (Kirtley, 2006). Baker (2006) identified four reasons for performing clinical gait analysis:

1. diagnosis between disease entities;
2. assessing the severity, extent or nature of a disease or injury;
3. monitoring progress in the presence or absence of intervention; and,
4. predicting the outcome of an intervention (or the absence of intervention).

While observational gait analysis is low cost, quick, and simple to perform, given the limitations of the unaided human eye, it suffers from the inability to observe small deviations in gait. Hence, a more accurate form of gait analysis was needed. The Vicon software coupled with an array of cameras, provides this accuracy.

Vicon and Gait Analysis

According to Baker (2006), modern clinical gait analysis can be traced back to the early 1980's, and while the technology itself has advanced at a significant rate, the concept has remained the same. Contemporary three-dimensional human motion capture such as the Vicon setup involves the use of high-speed cameras (typically ~250Hz) to track retro-reflective markers within a given volume, and as long as any two cameras can identify a single marker in space, its three dimensional position can be determined. The markers are placed at strategic anatomical locations (bony landmarks or prominences) on the body and motion is tracked and recorded in three dimensions in real-time. If the kinematic information obtained from the cameras is combined with kinetic force plate data, a comprehensive picture of the subject can be generated, particularly if oxygen consumption information and lower limb muscle activity is collected simultaneously (using electromyographic or EMG sensors to collect muscle activation information). For clinical gait analysis purposes, three-dimensional motion analysis provides important information needed to optimise the care of patients with complex walking

A traditional animation uses IK handles (which behave much like a muscle does in relation to a bone). These affect the rotations on joints which distort the skin over time generating the illusion of change and movement. Each major action of the character must be manually generated at key-framed intervals. Alternatively, human motion capture actions can be generated by actors with markers attached to them creating a data base of movements which can be used to drive a character rig.

The crucial difference between a traditional animation and that generated by human motion capture is the direction the action is processed by the computer. A traditional animation uses IK whereas motion capture uses Forward Kinematics (FK). For the case of the IK animation, the location and orientation of a rigged character is set at future points in time (or keyframes) and the computer interpolates the difference between the future keyframe and an earlier keyframe by working backwards. The computer does most of the between or 'tweening' calculations. By contrast, a motion capture FK animation must set the locations and orientations of all elements of a rigged character for every point in time (or frame). The computer is not required to do any interpolating between frames as the new locations and orientations are explicit. Thus, every frame is a keyframe.

Marker Placement

A typical rig used in a Maya animation is directed by the rotation values of joints about pivot points. The vertices of the mesh (the character's skin) are influenced by the transformations of the joints. Another type of mesh animation is surface deformation, such as in morphing and games animation meshes which do not use 'bones'.

Unlike a typical animated rig in Maya, the standard Vicon marker placements do not generate pivot-point locators. To overcome this, markers need to be strategically positioned such that their underlying pivot points can be extracted from the Vicon dataset post-capture (see Figure 2). As the goal of the motion capture data collection for this project was an animated skeleton in Maya, accurate skeletal pivot point locators (points in Maya holding location and translation values only) were needed. The authors recognised that to locate accurate pivot points in the human body markers were needed either side of the pivot points between bone junctions such that the centre point between the markers could be extracted from the data generated by the Vicon software. Markers need to be carefully positioned equidistant from a pivot point on the corresponding surface area of the body (see Figure 3).

Vicon to Maya

The Vicon software package supports data export in a number of different formats. The most useful for this project was the comma-delimited plain text CSV format.

The Vicon CSV dataset includes x, y and z values for each marker location and its corresponding location in time. This is represented as 3 columns (x, y and z) per marker location and a corresponding number of rows representing the time at which the location was recorded. For example, the data set for a 1 minute recording of a body with 40 markers filmed at 50 frames per second (fps) would include 3 x 40 columns by 50 x 60 rows plus 1 header row, 3.6×10^5 cells (this does not include the floor-embedded force plate data). However, the Vicon CSV file format cannot be imported directly into Maya. Therefore, a method for importing this data to a file format that Maya can recognise was needed.



Figure 3. Body with markers suitable for Maya pivot rigging (white markers are directly visible, red markers are included to indicate the corresponding marker that is masked or out of view). Note 3 of 10 Vicon Bonita cameras are visible in the background.

The Maya Embedded Language (MEL – similar to PERL) is the native scripting language used to communicate with the Maya programme. Python can also be used but it proved problematic hence the simpler MEL language was chosen. A MEL script was written to convert the Vicon CSV text file into coordinates which Maya translates as locators. The table of values exported by Vicon includes only the time-stamped x, y and z coordinate values for each tagged marker. Hence, the locators created in Maya by this script only have translation values – no rotational values. To create a skeletal rig for animation in Maya, rotation values are

required. Rotation values can be created from the locators using a number of different methods. A simple method is using vectors aligned between the locators.

Typically, any two points in space can be connected by a vector line which has two sets of coordinates, with three-dimensional data giving the vector line distance, orientation and direction. Orientation and direction can be used as a rotation value relevant to a world coordinate system. However, in Maya, bones or joints are rotated about only one end. Thus, although the vector line method is not directly applicable in this case the principle is helpful (Figure 4).

To generate a direction and orientation value for the locators in Maya they need to be 'aimed' at each other. The locators can be aimed at each other in order of their hierarchical chain, producing a series of rotations. However, a problem with this method is that when using Euler calculations in radians the locators can undergo gimble lock when rotating beyond 90 degrees (which occurs when an object is rotated beyond 180° causing it to flip in the opposite direction). This needs to be solved for each instance. A quick solve for this is to ensure that the overall axial orientation for all the locators is consistent. It is important that any two locators aimed at each other maintain a common orientation when one moves past the other. If not, gimble lock or flipping of axes may occur which will distort the skin of the rig when the animation is run.

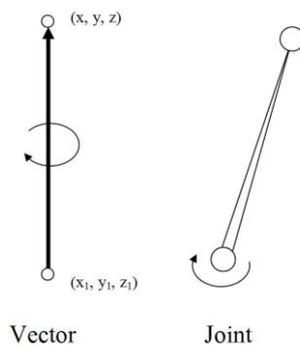


Figure 4. A vector has direction, location and orientation whereas a joint (bone) in 3D Maya only has location and its direction is driven by the rotation value of one of its ends.

Using the aim constraint function in Maya, locators are aimed at each other providing the rotation values necessary for the proper operation of the joint rigging. This is done in a hierarchical order from the root joint to the extremities. The rotation values, necessary for the joints to rotate and effect the skin mesh, is achieved using a combination of the aim, up and world vector parameters of the aim constraint in Maya.

MEL Scripts

To interpolate pivot point locations from the Vicon surface data, a script was written in MEL. It finds the

mid-point or centroid between any 2, 3 or 4 points given. This new centrepoint is then used as the pivot for rigging in Maya. However, before this function can be performed, the data from Vicon needs to be converted to a form readable by Maya.

Two MEL scripts were written to convert the Vicon data into something usable in Maya, a:

1. Vicon CSV plain text data table to Maya locator script; and,
2. centrepoint script to extrapolate a pivot point from two or more locators in Maya.

The CSV plain text to Maya script (CSV2MAYA) converts the Vicon data set into locators in Maya that include precise locations in three-dimensional space per time stamp (or keyframe). The centrepoint script is used to locate the pivot point between two or more locators representing data from markers on the surface of the body. For example, markers either side of the front and rear of the shoulder and equidistant from the pivot point between the humerus and the scapula at the bone junction are used to locate a third point.

CSV2Maya script

The CSV2MAYA script opens a csv file, identifies a start and end frame, and assigns the first line of text to a variable. It skips over the force-plate data (force-sensitive floor data) to the trajectory headings, which contain the body motion data and establishes header tags for the data in columns. If some data is missing, this is set to the arbitrary value $x=0$. The script reads in one line at a time, separates the data by commas and stores each value in an array. An x, y and z column is included for each group of three-dimensional data set. These data sets are related to individual marker positions with rows organised by timestamp or keyframe. When this data is imported and executed in Maya, via the MEL script editor, a locator cluster or point cloud is generated. The positions for individual markers are represented by the locators. Each locator is keyframed for the entire sequence. Individual keyframes may be missing from the dataset. These can be inserted manually between existing keyframes and edited in Maya's Graph Editor.

Centrepoint Script

The centrepoint script is executed on selected pairs, triads, or tetrads of opposing locators in the locator cluster already loaded in Maya. For each frame, the x, y and z values for each locator are retrieved. For bipolation, a simple distance calculation $(x_1 + x_2)/2$, $(y_1 + y_2)/2$, $(z_1 + z_2)/2$ is performed. A new locator is generated at the centrepoint between the two original locators. Triangulation and quadrilateration, follow a similar process:

$$(x_1 + x_2 + x_3)/3, (y_1 + y_2 + y_3)/3, (z_1 + z_2 + z_3)/3;$$

$$(x_1 + x_2 + x_3 + x_4)/4, (y_1 + y_2 + y_3 + y_4)/4, (z_1 + z_2 + z_3 + z_4)/4.$$

The new, centred, locators are used to establish the pivot points for the skeletal animation.

Inserting Joints between Locators

Once the Vicon data has been converted into a form that Maya can use to generate a locator cluster or point cloud, and the locators have been hierarchically aimed at each other, a rig can be constructed between the locator points. The rig is constructed using joints, much like the bone in a human skeleton. The joints (or bones) act as effectors for the skin mesh. They have ranges of deform effects which cause the skin mesh to animate with the joint's transforms. Maya's joint tool in conjunction with the snap tool is used to snap joints between locator points creating the overall 'rig'. The joints need to be parented to the aimed locators. This links the rotation values for the locators to the root of each joint. A skin mesh can then be bound to the overall joint rig which is deformed as the joints are rotated and translated. Figure 5 illustrates the sequential process.

Final Animation Result

When the animation is executed, the final skinned rig reanimates the original motion that was captured with the Vicon cameras. Multiple virtual cameras can be set up in Maya to view the animation from almost any angle. This provides insights not possible from the fixed cameras in the Vicon laboratory alone. Segments of the keyframed animation can be edited, manipulated and cycled to generate many different outcomes than the original capture. This provides a fertile source of human motion animation for use in other animation contexts – such as animating an animal to look like it has the attributes of a human, or a robot for the same reasons. Principally, it provides a source of complex human motion that would be difficult to animate manually.

Discussion

Unlike directly exporting the Vicon data to *MotionBuilder* or *Blade*, employing the process outlined in this project in conjunction with standard animation editing in Maya, students gain a fundamental understanding of the underlying principles involved. This provides critical learning opportunities that are useful

when they need to solve problems in the future. For example, when a marker is hidden from the cameras at some stage during a human motion capture in the Vicon laboratory, its position cannot be recorded for that period of time. This manifests as an empty value in the dataset. In Maya the locator for the missing data is set to the origin (x, y, z => 0, 0, 0) for a period of time. To correct this, the Maya Graph Editor can be used to isolate the frames either side of the missing data and new keyframes inserted. Keyframes are inserted for each of the x, y and z axes independently. As such, the operator can see exactly how each keyframe fits between the existing keyframe data. This corrects the missing keyframes in a manner that ensures a smooth animation. Critical here is that the student has the fundamental knowledge necessary to see what the problem is and how to correct it, because they have a better understanding of the overall concepts and processes involved. By contrast, when importing Vicon data directly into *MotionBuilder* or *Blade*, it is automatically converted to a fully formed skinned, animated, rig. And, if an error occurs it is not clear what type of error it is or how to correct it. The operator can manipulate the rig to overcome the missing data but the relocation of the missing keyframe data results in only an approximation of the desired outcome. Thus, when using *MotionBuilder* or *Blade*, incomplete Vicon datasets may instead need to be re-shot rather than solved as outlined in this project.

The students gain their understanding of the fundamental principles of human motion capture because the processes involved are open to scrutiny at every stage. From strategic location of markers on a body in the Vicon lab to using the CSV file, students can see how each marker on the subject is capture in time by its x y and z coordinates. Then, by selecting 2 or more opposing locators in Maya and executing the centrepoint script they see how a pivot point is extracted. Together with their prior knowledge of traditional animating rigging and skinning techniques they are able to make the necessary conceptual connections between a manually constructed rig and one generated from the Vicon data-point cloud. Moreover, they are able to leverage their prior knowledge and newly gained knowledge to identify and solve most problems as they arise (for examples, see <http://wyeldsoft.com/mocap/>).

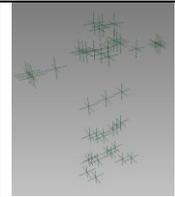
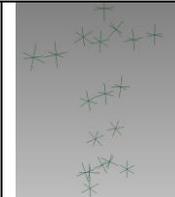
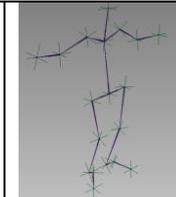
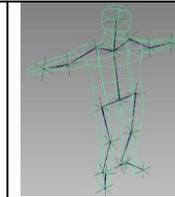
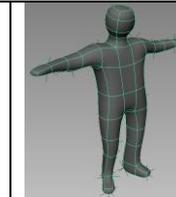
						
The original point cloud	With centred pivot points determined	Centred pivot points only	Locators hierarchically aimed	Joints snapped between aimed locators creating a rig	Polygonal mesh bound to the rig	The final flat-shaded output

Figure 5. The sequential process: from importing the Vicon data points, extracting pivot points, rigging and skinning

Conclusion

Undergraduate students of 3D media need to understand the fundamental underlying principles of human motion capture as it is used in the animation, VFX and the games industry. This is critical for them to be able to solve problems related to the use of motion capture at a fundamental level. There is little literature or research related to the procedures for generating motion capture and its use in 3D animation from first principles. This project goes some way to addressing this apparent literature knowledge gap in the field.

Although motion capture can be achieved using just two cameras the datasets generated are limited by the short angle view used. The University's Vicon laboratory provided a sophisticated facility for students to explore motion capture at a similar level to that used in the industry.

To achieve the goal of creating an animated rig in Maya from Vicon data, two MEL scripts were produced. One converts the Vicon data to a form that communicates with Maya and the other generates pivot data from surface data necessary to drive the rig's animation.

These scripts make it possible for students to generate motion capture data that can be used to produce smooth human motions assigned to rigs and meshes (skins) of their own design for use in animation, VFX or games. Moreover, as the scripts are written in plain text, their underlying functions and parameters are easily accessed and the connections between the information contained and outcomes better understood.

References

- [1] Adam, H. C., 2014, *Muybridge: The Human and Animal Locomotion Photographs*, Taschen.
- [2] Baker, R. 2006, "Gait analysis methods in rehabilitation", *Journal of NeuroEngineering and Rehabilitation*, 3, 4. <http://doi.org/10.1186/1743-0003-3-4>
- [3] Gage J.R., 1994, "The role of gait analysis in the treatment of cerebral palsy", *J Pediatr Orthop*, 14:701–2.
- [4] Kirtley, C., 2006, "Clinical Gait Analysis, Theory and Practice", Elsevier Health Sciences, Churchill Livingstone, ISBN: 0443100098
- [5] Korc, F., and Hlavac, V, 2008, "Detection and Tracking of Humans in Single View Sequences Using 2D Articulated Model", in Rosenhahn, B., Klette, R. and Metaxas, D., (eds) *Human Motion: Understanding, Modelling, Capture, and Animation*, Springer, pp. 105-132.
- [6] Molet, T., Boulic, R., Thalmann, D., 1996, "A Real-Time Anatomical Converter for Human Motion Capture", in Hegron G. and Boulic R, (eds.), *Computer Animation and Simulation'96*, Eurographics Animation Workshop, Springer-Verlag, pp. 79-94.
- [7] Muybridge, E., 1878, *The Horse in Motion*, source: Library of Congress Prints and Photographs Division;
- [8] <http://hdl.loc.gov/loc.pnp/cph.3a45870> [02-12-2015]
- [9] Silaghi, M-C., Plänklers, R., Boulic, R., Fua, P., and Thalmann, D., 1998, "Local and Global Skeleton Fitting Techniques for Optical Motion Capture", in Magnenat-Thalmann, N., and Thalmann, D., (eds.), *Modelling and Motion Capture Techniques for Virtual Environments*, Springer, LNCS, Vol 1537, pp26-40.