Development of an XML Web based motion capture data warehousing and translation system for collaborative animation projects

Carlos R. Morales
crmorales@tech.purdue.edu

Department of Computer Graphics, Purdue University, Knoy Hall Room 363, West Lafayette, Indiana, 47907, USA

Abstract

While motion capture has been hailed as a way to achieve extremely realistic animation in a cost-effective manner, many animators are reluctant to use it or to intermix it with keyframing techniques in situations where they are required to collaborate with other production personnel due to the lack of consistent software performance. The myriad of different file formats and the performance of those formats across different animation packages have limited the penetration of this technology into the production environment. The author details the development of an ASP web based application which archives segmented motion capture data into an intermediate XML format and then allows the users of the system to dynamically retrieve the data in standard Biovision segmented data format (BVA), Lightwave3D motion dump format, and Adaptive Optics Associates (AOA) motion capture format. The net result is a dynamic web enabled systems that allows users of Avid SoftImage 3D, AliasWavefront Maya, Newteck LightWave 3D, and Kinetix 3D Studio Max to exchange segmented motion capture data.

1. BENEFITS OF MOTION CAPTURE

While the use of motion capture in the production environment has always been received with mixed feelings, it is evident that technological advances in both the capture and application of motion to animation will continue to play a significant role in the industry. A quick survey of Hollywood blockbusters illustrates motion-capture being used by the larger production companies in movies like Titanic, The Matrix, and The Mummy. On television motion capture is used to generate real-time character animations, and in the gaming industry to create more realistic-game play. A stroll down the show floor of SIGGRAPH quickly reveals increasingly better motion capture products from companies like Polhemus, Ascension, and others.

While production companies have embraced motion capture as the answer to some specialized situations, such as the animation of crowds and the animation of central computer generated characters, the majority of work is still done though keyframing. The debate concerning keyframe animation versus motion capture revolves around the issues of control and speed. Animators want to use a method that gives them the most control while completing the work as quickly as possible.

The lack of an industry wide standard for archiving and exchanging motion capture data has contributed to the reluctance of animators to utilize this technology. If a company wants to make motion capture a central part of their production pipeline, they have been forced to either do all of their work within one package, or house a staff of technical directors who can write custom in-house utilities to bridge the gap among the different applications. The net result has been that only the larger production houses have been able to devote the resources to use motion capture in projects that require the use of multiple animation packages.

2. BENEFITS OF MOTION CAPTURE

Proponents for the wide usage of motion capture in the production environment see it as a cost-effective way of producing realistic animation. Emmanuel Javal from Medialab tells us that keyframe animation will give way to motion-capture as the prevalent mode for animation production, because it is a more cost-effective, production worthy alternative [Robert99B]. Another major argument is that motion-capture can create extremely realistic animation, which would be extremely difficult to achieve using keyframe techniques [Baumg99H]. For The Mummy, ILM used an optical system from Oxford Metrics to capture Arnold Vosloo's movements and map the motion to the main character. According to ILM's Jeff Light, motion capture allowed them to capture the essence of the actor's movement [Chron99E].

The ability to build reusable motion libraries is also a reason for using motion capture. "Once we have a motion captured, we can use it as many time as we want to," says Richard Fiore [Fiore99R]. High Voltage Software (HVS) used motion files captured for their NCAA Final Four video game and in the generation of their proposed Pacers animated opening. After the proper skeleton has been set-up in the animation package, motion libraries can be reapplied with just a few mouse clicks. The result is saved time and money. Dominique Pouliquen, marketing manager at Medialab tells us that motion capture can lower the cost of producing a 3D animated show to almost the level of doing a 2D show [Robert99B]. Once the initial set-up is performed, animation can be captured at an astonishing rate. HVS captured all of the motion for their NBA Inside Drive 2000 in two days [Fiore99R].

This proves extremely useful in situations where real-time rendering is necessary. Donna Coco tells us that real-time animation would be impossible without the use of motion capture [Coco97D]. On the show Canille Peluche, Mat the Ghost is captured, composited with live footage, and broadcast in real-time [Sturm94S]. Without the ability to capture the character's movement and render the results in real-time, it would be impossible to have the character react to actors in a live broadcast setting. In gaming scenarios, where the rendered graphics need to react to the player's movement, motion capture makes it possible to generate realistic animation. In Parasite Eve animators from Square Soft used motion capture for situations where "there is a lot of physical dynamics to the motion and you really see the gravity of the character…because that sort of movement can be really hard to achieve in keyframe animation."[Baumg99H]

3. DIFFICULTIES WITH MOTION CAPTURE

While the benefits of using motion capture are numerous, the lack of acceptable motion capture tools has limited its use among animators. Henry Baumgartner tells us "A particular difficulty comes in manipulating the movements further after they have been captured."[Mahon97D] Emilio Chronopolous, character animator NuFX, concurs, "Mo-cap is great if you don't have to change anything."[Chron99E] Because each frame is a keyframe, it makes any changes made by the animator result in jerky motion, if the animator does not rebuild the motion curves. It is often better to use the mo-cap as a template for generating new motion curves that can be changed and animated by hand [Mahon97D]. This process involves extracting keyframes for key poses in the animation and then keyframing the rest to approximate the original motion-capture data. In effect, the animator ends up spending more time cleaning up the motion capture data than actually animating.

Animating transitions for motion capture segments can also prove to be very time consuming. If the motion capture data is to be applied to in a game environment, animators have to hand animate moves to transition among motion capture data sets. In most situations the main animations are captured, but the transition animations cannot be anticipated because there are dependent on the actions of the player. These transitions can be extremely difficult because the animator has to match the realism of the two motion capture segments. In generating NBA Inside Drive 2000 and NCAA Final Four 1997, HVS animators spent a considerable amount of time animating transition moves [Fiore99R]. Tools which allow motion capture data sets to be smoothly transitioned, such as Character Studio, do exists but require the motion capture data set to come in using a very limited number of file formats.

A significant problem of using motion capture in the production environment stems from the lack of integration among motion capture hardware/software developers and animation package developers. Brennan McTernan from 3D Creations tells us "there really has not been a concerted effort by the 3D software manufacturers to find out how this stuff [motion capture] is used in production."[Doyle99D] Most high end animation software has really good support for their own native motion capture format and varying degrees of success with industry standardized formats such as Biovision (BVA and BVH) and Acclaim. In SoftImage 3D the straight application of motion capture files to articulated skeletons often results in flipped chains [Jeffr99E]. Maya requires the use of a third party MEL script to accept BVH or Acclaim [Alias99W]. Flipped chains are also common (see Figure 1). BVA, a positional motion format, is unsupported without custom scripting (See Table 1). Discreet's 3DS Max requires an additional plug-in called Character Studio to read either format [Discr99D]. Newtek's LightWave uses its own proprietary format [Newte98L] (see Table 2).

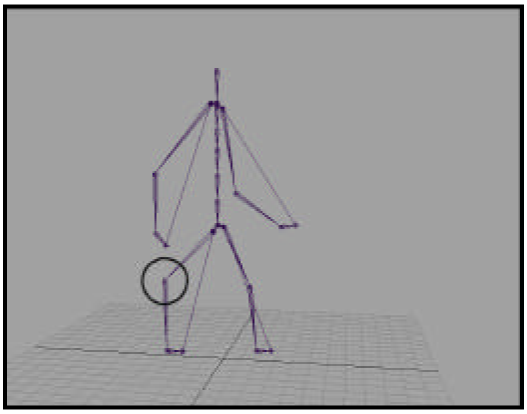

Figure 1. Flipped Articulated Chain

0.0251	0.57	-10.0837	0	0	0	1	1	1
0.0332557	0.517897	-10.0526	0	0	0	1	1	1
0.034022	0.515758	-10.02118	0	0	0	1	1	1

Table 1. LightWave 3D Proprietary Motion File

Segment: Chest
Frames: 3
Frame Time: 0.033333

XTRAN	YTRAN	ZTRAN	XROT	YROT	ZROT	XSCALE	YSCALE	ZSCALE
INCHES	INCHES	INCHES	DEGREES		DEGREES		DEGREES	
0.0225	0.57	10.0837	0	0	0	1	1	1
0.0332557	0.517897	10.0526	0	0	0	1	1	1
0.034022	0.515758	10.02118	0	0	0	1	1	1
0.0347978	0.513586	1.989462	0	0	0	1	1	1

Table 2. Biovision BVA Motion File

The result is that the same motion-capture files react differently in these packages. The problem, Mahoney says, is that most applications treat motion capture as a plug-in or as something extra and not as an integral piece of the animation process [Mahon97D]. The result is a tremendous amount of clean up and fidgeting on the side of the animator to get the program to work properly with the supplied motion files. It also locks the animator to a specific program once he/she has decided to use motion capture, because of the unpredictability of the motion-capture data in other programs.

The program's operation for manipulating motion-capture also varies. LightWave provides no specific motion capture features other than assigning nulls [Newte98L]. Maya and SoftImage have a more complete set. Maya has the ability of letting animators and technical directors write their own code [Alias99W]. Character Studio shines as it lets the user blend, concatenate, and filter motion [Discr99D]. Silva et al. identified these operations as critical for an effective motion capture system [SILVA97]. Brian Schultz, animator at HVS, says that Character Studio gives him more options for directly manipulating motion capture data than he has found in Maya or LightWave [Newte98L]

4. A CLOSER LOOK AT APPLYING MOTION CAPTURE DATA

While the application of motion capture data varies greatly from program to program, the actual capture process is relatively the same from the practitioner's point of view. Regardless of the technology used, most manufacturers encapsulate the process to generate the necessary motion files. Most producers are insulated from the actual capture process and only need to concern themselves with the implications and limitations of the technology they select. Many production houses subcontract the capture process to companies that specialize in capturing motion. The firm that actually performs the capture process will deliver the correct files. HVS used hired House of Moves to capture data for its NBA Inside Drive 2000 [Fiore99R]. Digital Domain had House of Moves do the capture for Titanic [Titan98C].

The selection of which motion-capture technology to utilize is largely dependent on the purpose of the data. Two main questions have to be answered. First, does the captured data need to be applied in real-time, and second, does the project require rotational and positional tracking. For production work, rotational data is extremely important because it allows the data to be used with a variety of different sized articulated skeletons [SILVA97].

Tethered solutions can capture rotational and positional data in real-time. The subject wears mechanical and/or magnetic sensors that are wired to a computer for real-time calculation [Coco97D]. The benefit with these systems is that the data is available for real-time mapping and does not require post processing. This can be a big benefit for trade-show displays or shows like Canille Peluche. The negative side is that the wires and/or mechanical sensors can get in the way of the performance. Another negative is that most systems using magnetic sensors require the use of special flooring to prevent magnetic interference from metallic support beams used in most commercial buildings. For NBA Inside Drive 2000, HVS decided a tethered system would not work for capturing basketball movements [Fiore99R].

In an optical system the user wears a series of reflective balls, which are captured from a variety of different cameras. Using triangulation techniques the computer can determine the location of each sensor [Sturm94S]. Because the footage must be analysed, the capture session must be processed off line. As a result most optical systems cannot be used for applications that require real-time application of the captured data. Each dot looks exactly the same to the system; the processing software must keep track of them and calculate rotational data based on the position it has calculated [Coco97D]. A benefit of these systems is that they do not hamper the performance of the actor(s) because the actors are not required to wear tethered suits. The added time for processing the frames off line is often considered a negative aspect of optical solutions. Another strike against optical motion capture is the fact that sensors can become optically occluded during the capture session.

Products using any variety of capture techniques can deliver hierarchical or point independent data. Hierarchical data such as that supplied in Biovision's BVH or the Acclaim format describes motion in terms of rotations based on a pre-defined skeleton. Thus, the actual motion capture data file describes the skeleton for which the data was captured. In the case of the BVH format the skeleton is described within the file [Biov99I]. In the acclaim format the motion-file is delivered in two files, a skeleton file and a motion file [Accla99I]. Both are needed for the motion to work. If the program reading the file understands hierarchical data, it can be a tremendous time saver. HVS used Acclaim for NBS Inside Drive 2000 [Fiore99R]. Non-hierarchical data such as BVA describes only the location of points. This makes this format capable of describing just about any motion, but it can require extra set-up time to use in character animation. If the animator wants to use the data to drive character animation, he must define a skeleton and then assign the position and/or rotational data to the skeleton based on the position of the data points in the file. This often results in flipped chains.

5. DEVELOPMENT

The author decided to construct an application that would allow animators to transparently share segmented motion capture data.

In order for the system to be successful in a production environment, it would need to offer immediate translation, indexing, and archiving of motion capture files in a network environment. This would allow multiple animators working on the same project to engage in live on-line collaboration.

The author elected to use Microsoft's Active Server Pages (ASP) technology to construct the application because of its extensibility and ability to interface with databases. ASP allows content creators to generate dynamic web content by relying on server-side components. By using ASP, the author was assured the interface into the application would be cross-platform.

With the underlying technology selected, the author turned his attention towards devising a method for translating the motion capture data files. The author restricted the system to only work with segmented data files, which describe motion based on positional data. Hierarchical motion files that describe motion based on a predefined skeletal structure, defined at time of the motion capture session, were place outside of the scope of this project by the author. This restriction would decease the complexity of the tool. Positional data can be used in a hierarchical system by constraining the end of bones effectors in the 3D animation package to the nulls created by the positional data.

Three files formats were selected due to their ability to be imported into almost all of the major 3D animation packages. Biovision segmented format (BVA) was selected because it imports into SoftImage 3D, AliasWavefront Power Animator & Maya, Pixel's Pixel 3D, and Kinetix's 3DStudio Max [requires free plug-in]. LightWave 3D motion dump format imports into Newtek's LightWave, and Adaptive Optics Associates (AOA) format allows users to import into programs that read raw Cartesian positional data.

In deciding how to translate the files, the author turned to XML because it would allow the application to separate the content of the motion capture file from the presentation of that data to the programs using it. This became a necessity because the different programs would behave differently when supplied with the same motion data file. By parsing the data provided by the user into an XML document and then developing a set of schemas for each viable output, the application would be able to use a single content file, which could serve multiple animation applications. The consolidation of the data into a single content file ensured consistency across applications. Regardless of which application an animator used in a production environment, he/she would get the same result when accessing the same file. Changes to the single content file would be propagated and reflected to all users of that file regardless of its form of presentation. If additional formats were needed at a later date, the author would only need to add a new schema that would allow the system to serve the original content file in the new format.

The author decided to house the master XML file and its supporting schema files within a Microsoft Access database that would be accessed by the ASP application. The ASP application would deliver the file to the user in the requested format. This approach would require less maintenance than having the ASP application serve out XML and then write plug-ins for each of the desired 3D animation packages to read in the XML file.

Before the ASP application could be used to parse out in-coming motion-capture files into XML, the author had to develop a structure that would hold the necessary data to reconstruct any of the input files. Upon examining the content of Biovision Segmented files, LightWave 3D motion dump files, and Adaptive Optics Associate files, the author concluded he would have to track the following properties: number of nodes, number of frames, length of time between samples, x position, y position, z position, and the name of each node. This analysis led to the XML structure revealed in Figure 2.

```
- <Mocap>
  - <General_Information>
      <NumberNodes />
      <NumberFrames />
       <SampleFrequency
         />
    </General_Information>
  - <Nodes>
    - <NodeName>
        <XTrans />
        <YTrans />
        <ZTrans />
        <XRot />
        <YRot />
        <ZRot />
        <XScale />
        <YScale />
        <ZScale />
      </NodeName>
    </Nodes>
  </Mocap>
```

Figure 2. Derived XML Motion Capture Structure

Closer examination of files revealed the author would have to decide on an objective standard that all of the files would be translated into, even when the files are of the same type. The same motion capture file behaves differently when imported into animation packages with different coordinate systems. For example, reading the exact same BVA file into Alias Power Animator and SoftImage 3D will result in different results because the coordinate system in Alias Power Animator is Z-axis up and SoftImage 3D is Y-Axis up. If the application were to truly allow collaboration among animators, the author would have to make note of these differences. Differences in scale would also have to be tracked.

The author developed a schema that the ASP application could use to parse and translate each of the file formats. The BVA format required three schemas to account for differences in behaviour in Power Animator, 3D Studio Max, and SoftImage 3D. Pixel3D interpreted the file identically to SoftImage 3D. Thus, the same schema could be employed. Microsoft's XML parser was used to parse the data.

The application provides the user with the ability to archive the data into the company's database when translation is requested. This enables the company to build up a library of motion capture moves as a by-product of the users translation motion files. The more files translated, the larger the motion capture database grows. To allow future users of the system to search the database without having to open each XML file, the system keeps the description of the files in a separate field within the Microsoft Access Database. Figure 3 details the all of the components of the system.

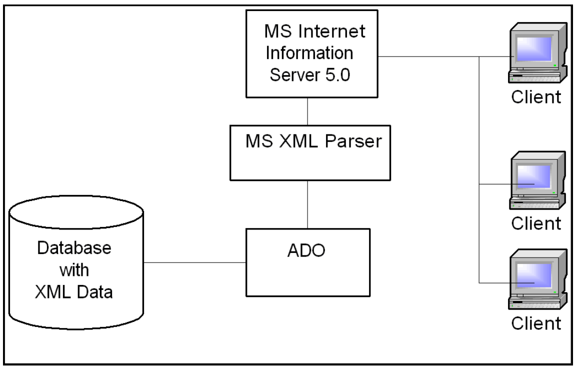

Figure 3. System Overview

6. CONCLUSION

Overall this application alleviates the process of using motion capture data in collaborative animation environment by providing users with a means to consolidate their motion data into a single data format which can be correctly accessed by the entire production team in an immediate networked searchable environment. The combination of XML and ASP facilitated the development.

Improvements to this application can be made in obvious areas. One, a component could be written in each of the client 3D animation packages that would allow searching, publishing, and retrieval of the data from within the native environment of the 3D program. For example, a plug-in could be written in Max Script that would access the web server and transparently download the latest version of the requested motion capture file. It could also allow the user to search the database without leaving 3D Studio Max. A second area in which this application could be improved could be extending the native parsing capabilities of the system to allow the user to read in and parse hierarchical data such as Biovision Hierarchical Files (BVH). This would allow users to work with files with predefined skeleton without having to spend the time needed to set up a constraint system. The net result would be two to three hours of saved time on the animators end when exchanging motion data requiring a hierarchy. In the end, the system does make the process of collaborating on projects with motion capture data easier.

7. REFERENCES

[Accla99I] Acclaim Incorporated, Acclaim ASF/AMC File Specifications. http://www.acclaim.com, 1999.

[Alias99W] AliasWavefront. Maya 2.0 Users Manual. Toronto, Canada, 1999.

[Baumg99H] Baumgartner, Henry. How to Catch a Ghost. *Mechanical Engineering*, Vol. 121, p108, April 99

[Biov99I] Biovision Incorporated, Biovision Motion Data Specifications. http://www.biovision.com, 1999.

[Chron99E] Chronopolus, Emilio. Personal Interview. *Character Animator*. NuFX, Rolling Meadows: Illinois, 1999.

[Coco97D] Coco, Donna. Motion Capture Advances. *Computer Graphics World*, Vol. 20, Issue 11, p37, Nov97.

[Discr99D] Discreet. 3D Studio 3.0 Users Manual. San Rafael, CA, 1999.

[Doyle99D] Doyle, Audrey, *Computer Graphics World*, Vol. 22, Issue1, p58, Jan99.

[Fiore99R] Fiore, Richard. Personal Interview. *Art Director*. High Voltage Software, Hoffman Estates: Illinois, 1999.

[Jeffr99E] Jeffrey, E. Personal Interview. *Character Animator at High Voltage Software.* Hoffman Estates: Illinois, 1999.

[Mahon97D] Mahoney, Diana. Face Tracking. *Computer Graphics World*, Vol. 20, Issue 4, p23, Apr97.

[Newte98L] Newtek. LightWave 3D 5.5 Users Manual. Austin, Texas: Newtek, 1998.

[Robert99B] Robertson, Barbara. Bad to the Bone. *Computer Graphics World*., Vol. 22, No. 5, p34 May 99.

[SILVA97] SILVA, F., et. al., An Architecture for Motion Capture Based Animation. *In Proceedings of SIBGRAPI'97, X Brazilian Symposium of Computer Graphics and Image Processing*, pages 49-56, oct. 1997.

[Sturm94S] Sturman, David, A Brief History of Motion Capture for Computer Character Animation. SIGGRAPH 94, 1994.

[Titan98C] Titanic Special Issue, CineFex, No. 72, 1998.