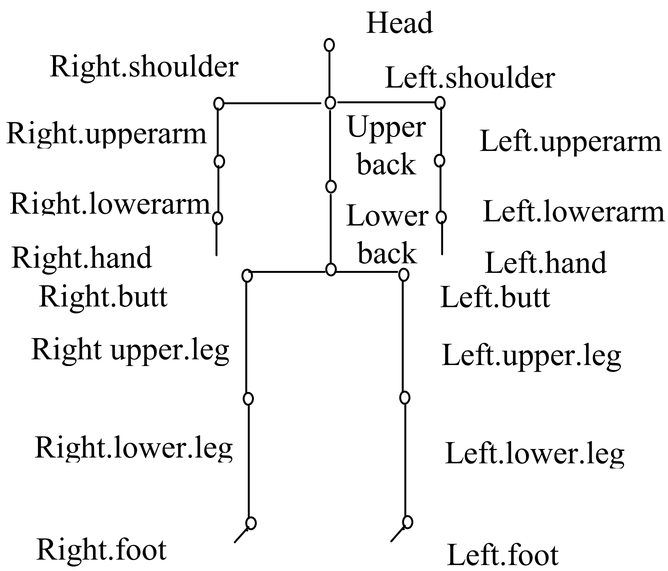
Character Animation Using Motion Capture

What is Character Animation?
Character animation can be defined as the expression of emotion or behavior of a live or inanimate object through the use of motion.

How is A Character Modeled?
To represent the motion of a living creature such as a human, animal or arthropod (insect), it is convenient to store the relationships between each movable part of the creature eg. Head is attached to the neck, left foot is attached to the lower left leg and so on. Each movable part is referred to as a bone, the attachments between each bone are called joints and the entire collection of bones is referred to as a skeleton. For example, a basic mammal (without fingers or toes) would take around 20 bones:



A skeleton can be defined as an array of bones. Each bone is defined in terms of a length, a default direction vector, a name and the name of the parent.

The parent of a bone is the bone which is attached to that bone and will cause that bone to move if it moves. eg. If you turn your neck, your head will also turn. The only bone which does not have a parent is the root node. It is not really a bone in a physical sense but rather it is used as a convenient way of moving the entire skeleton with a single translation or rotation. So, for the example above, the list would be as follows:

Bone	Parent	Length	Direction
Root	---	-	---
Head	Upperback	3	010
Upperback	Lowerback	5	010
Lowerback	Root	5	010
Left.shoulder	Upperback	8	100
Left.upperarm	Left.shoulder	3	0-10
Left.lowerarm	Left.upperarm	3	0-10
Left.hand	Left.lowerarm	2	0-10
Right.shoulder	Upperback	8	-100

Right.upperarm	Right.shoulder	3	0-10
Right.lowerarm	Right.upperarm	3	0-10
Right.hand	Right.lowerarm	2	0-10
Left.butt	Root	3	100
Left.upperleg	Left.butt	5	100
Left.lowerleg	Left.upperleg	4	100
Left.foot	Left.lowerleg	2	001
Right.butt	Root	3	-100
Right.upperleg	Right.butt	5	100
Right.lowerleg	Right.upperleg	4	100
Right.foot	Right.lowerleg	2	001

Skeleton in Motion

Since each bone has a direction vector and a length, it is possible to use rotation matrices to transform each bone to a new position. Generation of the rotation matrix can be performed by the mathematical concatenation the XYZ rotation matrices. The new position of a skeleton after motion can be evaluated fairly easily. Since every bone has the following set of local information:

Start-point - The end attached to the parent
End-point - The end where children are attached
Relative matrix - Matrix generated from local rotation/translation values
Absolute matrix - Matrix generated from local matrix and parent matrix

The first stage performed is to evaluate the root node. This will involve setting the start-point, end-point and the "relative" (R) matrix. Since this is the root node, the "absolute" (A) matrix is identical to this.

The second stage involves the generation of the A-matrix for each bone in the model. For each bone in the model, the following steps are performed:

[1] The parent bone is determined. Only if this bone has already been evaluated, can the following steps can be performed. The ID of the parent bone can be represented either by a name-tag or by an index value.

[2] The end-point of the parent bone is used as the start-point of the current bone.

[3] The R-matrix is calculated (if this has not already been done).

[4] The R-matrix is combined with the parent's A-matrix to generate a new A-matrix.

[5] The direction vector is combined with the A-matrix to determine the current orientation of the bone.

[6] The current orientation of the bone is added to the start-point in order to generate the end-point.

This process is repeated until all bones have been evaluated.

Animation of Motion
Traditionally, motion for animation has been created by specifying the position of a character at each instant in time (each of them is called a keyframe). In a method similar to pencil drawn flip-books, this can give the appearance of motion through the continuous presentation of different positions of a character. For example, we may animate a simple walk cycle with 8 keyframes. The first four frames will have the left leg and right arm swinging forward and backwards to rest, then the last four frames will be when the right leg and left arm swing forward, and the cycle will continue from the beginning again.

It has the drawback of being laborious as well as requiring a great deal of skill to create convincing motion by specifying a series of individual poses as properties of the motion are created over many individual poses. While computers can reduce some of the labor by automatically interpolating between keyframes, it still requires talent and training. It is particularly difficult to create motions that are realistic and/or accurately mimic subtle characteristics, such as a particular person.

Another method is to use algorithm or simulation methods to generate motions based on descriptions of goals. While such methods have the promise of generating motions for non-experts by allowing them to simply specify their needs, they are, at present, of limited use, as there has been no systematic way provided to create new behaviors. One key problem facing algorithmic methods is how to describe a complicated motion.

For animation which requires a large amount of movement (say dancers on a stage), motion capture can be used. Techniques of motion capture has been studied since the beginning of the 80's. Nowadays, there are four main methods of motion capture, namely mechanical, magnetic, optical and hybrid method.

Mechanical systems
It composed of potentiometers (or sliders) that measure the position or orientation of joints in an object. Its similarity with conventional stop-motion techniques, that are widely used in movie production, allows a natural migration of traditional animators, thus increasing the popularity of this techniques. However, the realism of mechanically captured motions still depends, in great part, on the ability and patience of the animator.

Magnetic
Another type of motion-capture device is the electromagnetic tracker, which consists of a series of receivers or sensors, a transmitter and a control unit. The sensors are connected to the control unit and are placed on the body of the performer. The transmitter generates a low-frequency magnetic field, and as the receivers move through it, the control unit is able to track their signal in order to calculate their position in space. These systems have been widely used for many years in military applications and have a higher margin of error and smaller capture frequency than the optical types, but they are less expensive and capable of real-time feedback. These systems' biggest drawback is their susceptibility to interference by metallic objects. Depending on the type of electromagnetic field used (AC or DC), the tracker will be affected by different types of metals. This means that all capture sessions must happen in a controlled metal-free environment.

Optical systems
It is based on high contrast video imaging of retro-reflective markers, that are placed on objects whose motion is being recorded. This technique provides high sampling rates, but the recorded motion data must be post-processed using computer vision tracking techniques.

These systems were designed for applications are the most expensive, but they have the smallest margin of error and the highest capture frequency. Their main drawback is that most of these systems are not capable of true real-time capture, which is highly desirable in the visual effects industry. Today's state-of-the art systems have dozens of cameras and can capture hundreds of markers at frequencies of 1,000 samples per second or more.

In the tracking process, the controids of markers are matched in images from pairs of cameras, using a triangulation to compute the positional data of these markers in 3D space. This process introduces artifacts (offsets) into the final data. Some disadvantages of the optical process are the occlusion of one or more markers during the capturing session, the lack of angular data, and the sensitivity to background light and reflective objects.

Hybrid Systems
It combine both magnetic and optical technologies are being developed, but are not yet commercially available.

Steps in Creating Animation From Motion Capture
In this article, I will describe steps in creating animation from an optical motion capture system. The motion capture process consists of 4 major stages:

- In a typical motion capture session, the performance of standardized "gym" motion is first recorded for the skeleton fitting post-processing.
- Then, according to a work plan, multiple takes of each motion are rehearsed until the artistic director is satisfied with the performance of the artist. It is important to remember than only 2D marker images are captured using dedicated hardware. So the artistic direction aesthetic evaluation can only focus on the real performer's motion and not on the corresponding virtual character's motion.
- The post-processing phase currently occurs after the motion capture session and can last days or weeks depending on the complexity of the recorded material. The first stage is motion tracking, to identify the performer's motion expressed in joint angle trajectories.
- The last stage is the anatomic conversion producing the virtual character's motion while retaining the emotional subtleties conveyed by the real performer's motion.