UNIT – 4 VISUAL EFFECTS TECHNIQUES

What Is VFX?

Visual effects are special techniques used to create non-existent objects, fictional creatures, imaginary landscapes, and even bone-rattling actions negating the universal laws of physics. They are usually combined with live-action footage to produce realistic video content. In old films, imaginary scenes were made with the help of tricky physical methods like a multiple exposure or rear projection (used in 2001: A Space Odyssey, for example). However, the digital revolution boiled it down to computer-generated imaginary integrated into footage in the post-production stage.

VFX production is brought into play when actual shooting on a filming location is too expensive, dangerous, time-consuming, or even impossible. Yes, you can shoot a real tornado if you don't mind chasing one while risking your life. And you can build a copy of Titanic if you are James Cameron with the initial budget of over $ 100 million. But you can have both tornado and Titanic created with computer graphics and then embedded in your video. And that is the key aspect behind VFX – it allows for the implementation of creative ideas in the most suitable way to grab the modern audience's attention. And today's audience is rather visually demanding of any type of video content, be it a music clip, a promotional video, or a feature film.

Yet, VFX services aren't cheap, though they do save money in many scenarios. The costs may range from $25,000 to $35,000 for a 5-minute video, depending on the scope of work. Hollywood producersinclude tens of millions in the VFX budgets, with dozens of studios working on one project. For example, Avengers: Infinity War was created by 15 studios, with the total number of VFX artists reaching 2,080 people. So, many dedicated specialists are involved in the process of creating visual effects, and you will understand which and why as you will proceed with the article.

Visual Effects vs. Special Effects

In the first place, let's sort out where video visual effects start and special effects end. And no, they are not the same, though they often go hand in hand. Here is a simple example illustrating the difference. Say, you need rain in your video. You have several options:

to wait for a rainy day – no effects;

to create rain with the help of a hose – special effects;

to ask a visual effects artist for the rendering of rain in computer graphics – visual effects;

to use a hose on a production set and add computer-generated clouds to the footage later for a more realistic feel – SFX+VFX.

So, SFX is filmed right on set with real stage props or actors. VFX is a combination of real-life footage with added CGI.

Visual Effects Process and Techniques

Well, VFX is mostly associated with post-production, but it doesn't mean you need to think of it only after filming is completed. Quite the contrary – you should start choosing a VFX outsourcing service provider right after your project is green-lighted. If you fail to do it, you are likely to face many problems with incorporating visual effects into your video coupled with time-wasting and going over your budget. That's because the process of creating visual effects accompanies the whole video production flow, from the preparatory stage to the final result. We have collected seven techniques showing how VFX is integrated into the different stages and what outcomes you can get from them.

Previsualization

Making a storyboard is an old good method to get prepared for shooting. Nevertheless, computer-generated simulation of scenes provides even more opportunities to convert ideas into realization without a hitch. It is especially true for complex scenes with SFX and VFX, which require careful planning. Previsualization allows experimenting with staging, lighting, camera movement, and other aspects at the pre-production stage to reduce costs and time expenditures while ensuring the best possible result. Some of the finest pre-visualized sequences were created for Avatar, Life of Pi, and World War Z.

Matte Painting

This technology serves to combine live-action footage with a painted environment when it is impossible to shoot real-life objects or too costly to recreate them through set construction.VFX graphics help depict a castle of elves or landscapes of Mars using 2D or 3D digital tools. Modern matte painting allows for any manipulations with background images and any combination of shooting on location with computer graphics. It is perfectly demonstrated by the example of Martin Scorsese's Casino, where it was necessary to reproduce the neon street lighting of Las Vegas in the 1970s.

Compositing

When there are actors filmed on set and a computer-generated background from the previous passage, how are they blended together for a viewer to perceive the whole scene as real? Here is where compositingsteps forward to glue multiple layers together and deliver a seamless, realistic image that, in fact, has never existed. A compositing artist from a VFX production house will use various technologies like chroma key, camera tracking, or match moving to ensure the perfect consistency of final visuals, similar to that we all enjoy in Star Wars, The Lord of the Rings, or Harry Potter movies.

Animation

Any object that is supposed to move somehow in your video needs to be animated (if it is not a live creature, of course). Animation is a diverse world of techniques, starting from motion graphics used to add movement to text or abstract shapes and ending with motion capture for creating highly realistic computer-generated images. You can get snow or explosion thanks to particle VFX simulation or insert a 3D dragon into your footage – after the creature is digitally modeled, rigged, skinned, textured, and set in motion to look just like in How to Train Your Dragon.

Chroma Key

It is a famous technique of combining several images where an object filmed is located on a monochromatic background replaced by another image in post-production. This term is also used to refer to the background itself, which is usually green or blue. The choice is explained by the fact that these colors don’t match human skin tones. However, the background can be red or yellow as well, depending on the task set by the director, the equipment characteristics, and the colors of clothes and objects to be caught on camera. Objects are separated from a green or blue screen during the process called keying, with the results seen in Ed Sheeran and Justin Bieber’s I Don't Care.

Bullet Time Effect

Hello, The Matrix! The now-iconic scenes with “frozen” bullets gave the name to this technique for creating the illusion of transformed time and space. A VFX studio can create the effect of the camera and the viewer hovering around an object that seems to stop or slow down while actually rapidly moving. It doesn't need to be a bullet but the moment of a jump or fall, for example.The effect can be obtained by replacing a video camera with several dozen still cameras following the trajectory of planned shooting and firing sequentially.The resulting set of single shots is combined into a video sequence, creating the illusion of continuous movement.

Transitions

While moving from one scene to another can be implemented through a trivial cut, transition effects are great for adding spice to your VFX motion graphics, a music clip, or any other video content. Transitions have the power to create a certain mood, bring in dynamics, or control the pace of your film – everything depends on the director's idea and VFX artists' skills. The result of their work is shown through the Do What You Can promo for Samsung crafted by Frender.

Visual Effects Examples

To better understand some of the techniques described above, see how they are implemented in the Planet of the Apes series:
Over 1,500 shots were pre-visualized for Rise.
Matte painting allowed adding a computer-generated model of the Golden Gate Bridge to a physical set.
The apes, horses, and other minor objects were created digitally through motion capture and hand animation.
The fur and snow required plenty of particle simulation.
Surely, it is not the whole story, but we hope you've got the idea behind visual effects production.

Green Screen, Keying, and Rotoscopy

Green screen removal, keying, and rotoscoping are crucial techniques in the world of visual effects (VFX). These techniques form the backbone of most VFX shots and are used to create seamless and believable footage. Keying involves removing a specific color (mostly green or blue) from a shot and replacing it with another image or video. Green screen removal is a type of keying that involves filming actors in front of a green screen, which is later replaced with a background or other images. Rotoscoping is the process of manually tracing an object in a shot frame-by-frame to create a matte, which can isolate and manipulate the object in the scene.
Our experts at Motion Edits use green screen removal, keying, and rotoscopy techniques to create stunning VFX shots for their clients. These techniques allow them to manipulate the footage in ways that would be impossible with practical effects or in-camera techniques. These techniques provide a

high degree of control over the final image, allowing for the creation of complex and highly realistic VFX shots.

1. MOTION CAPTURE

What is Motion Capture, and How Does it Work?

Motion capture (mocap) technology is used to record movements and apply them to a 3D model. Physical mocap suits, specialty cameras, and advanced software are used to create photorealistic animations that can be used in film, sports, and even healthcare.

What is motion capture?

Motion capture records movement and translates it into data that can be read by animation software and applied to a 3D rig or character. It's a common misconception that motion capture projects need a big budget and an entire production team. With evolving technology, you can even use your cellphone to do basic motion capture. For example, Instagram filters use a type of mocap to track your face in real-time and apply simple animation overlays. It's a similar technology that is used in Rokoko's facial mocap app for iOS, which allows 3D artists to capture facial motions as blendshapes to apply them on their custom characters in their 3D animation projects.

The rise of motion capture as a Hollywood must-have

Motion capture was first used in the film Sinbad: Beyond the Veil of Mists (2000). In the next few years, it was quickly popularized by the then-revolutionary animated character Gollum in The Lord of the Rings. The animated character interacted realistically with his live-action co-stars by relying on the mocap technique developed by Weta studios. Keep reading until the end of the article to see a video of what that looked like in practice.

From then on, mocap has become an almost mandatory feature in major films requiring VFX. In the late 2010's motion capture evolved from tracking humans to tracking animals. It's now possible to record movement from popular domestic animals such as dogs and horses.

What applications are suitable for motion capture?

Motion capture can be used for many types of projects, not just VFX. These include:
Game animations use mocap to quickly build up a vast library of motions for each game character.

Previs (also known as previsualization) happens during pre-production. It's when the creators bring the static storyboard to life. In complex scenes, directors will often use mocap to block out the motions of the scene and more accurately prepare for shooting and VFX.
Humanoid fantasy characters need to move realistically to avoid the uncanny valley effect. And that's what mocap helps animators achieve.
The health and sports industry is a big user of motion capture technology. It's been used to do everything from optimizing an athlete's tennis swing to injury diagnosis and rehabilitation.
In the military, mocap is used to create advanced simulations and improve training programs.

What to consider when starting a motion capture project?

First, you need to be aware of the four main types of motion capture, what they mean, and how to choose one that best fits your project needs.

Optical-Passive: Retroreflective suit markers & infrared cameras

Retroreflective markers are placed on actors via a tight-fitting suit and tracked via infrared cameras. Historically, this was the most common way of doing motion capture. Large studios commonly use this type as it's the most accurate, yielding the impressive photorealistic tracking required for feature films. However, it can be resource-hungry and isn't suitable to run on entry-level systems.

Optical-Active: LED suit markers & cameras

Light-emitting LED markets are placed on actors the same way as optical-passive tracking, and special cameras record their movement. This isn't used often anymore as the actors also need to carry some kind of charger or battery case, and the LED light can potentially spill into other filmed elements.

Video (Markerless): A sophisticated camera stage is used

Actors do not use suit markers of any kind. Instead, the acting area is covered by a grid on the floor and a network of cameras that shoot the scene from every possible angle. The recorded footage is analyzed by software and translated into motion data that animation software can read. However, the end result takes more time and includes more errors than other methods, meaning that a lot of data cleanup is needed in post-production. This type of motion capture is useful for large-scale productions that have post-production budgets. It captures the scene from every angle, reducing the need for retakes.

Inertial (Cameraless): Motion sensor suit

Unlike the other types, Inertial requires no cameras to capture the motion. Instead, inertial sensors (IMUs) are placed within a bodysuit and worn by the actor. The motion data is transmitted wirelessly to a nearby device. The gyroscopic motion sensors record the angle, position, and momentum of your body and accurately transcribe it into animated movement. This is the most cost-effective option and is popular with indie studios and game developers, like indie game developer Brian Parnell that completed all the character animations for his game "Praey for the Gods" with the Smartsuit Pro, get the full story here.

The benefits of using motion capture

There are three core benefits to using motion capture in your production

1. VFX costs and animation timelines are significantly reduced

Typically, 3D animators place keyframes for every major movement. Then, they adjust every frame with micro-movements. Sometimes repeating this process hundreds of times for each limb. Considering there are at least 24 frames per second, many productions quickly go over budget and miss deadlines due to animation. Using motion capture, the bulk of animation work is completed with the live-action actor's movements.

2. Facial animation becomes way easier

Accurate facial animation is known to be one of the most challenging tasks - especially if you're aiming for a photorealistic outcome (just check out the length of this 2021 paper!). With a simple mocap setup, you can capture basic facial animations. To capture the more realistic animations seen in Hollywood films, you'll need a more sophisticated setup that often includes a direct 3D scan of the actor's face to map movements correctly.

3. Previs for animations is cheap enough for small productions

Previs is essentially the previsualization of any movie, game scene, music video, or short film. It's often done through hand-drawn storyboards that are timed to voice-over or music. In productions that require a high level of planning (e.g. an animated dancer in a music video), you can use mocap to guarantee that your choreography is on time and in the frame. 3D Artist and Rokoko user Don Allen III is behind the motion capture recordings of Lil Nas X's Panini music video:

The 4 basic steps when getting started with motion capture

Exactly what you need to know if you want to use motion capture for your production

1. Decide on a type of motion capture

As you learned earlier in this article, there are four main types of motion capture. However, two types currently dominate the market; Optical-Passive and Inertial. We recommend that you only consider Optical Passive for larger projects with bigger budgets. It requires a bodysuit, software, and cameras capable of infrared capture. For all other purposes, the Inertial type is best suited. Intertial mocap can be captured in any location with any props as long as you have Wi-Fi access.

2. Decide on a system & software

Most motion capture systems provide their own propriety software built to perform optimally with their suit and/or cameras.

3. Make sure your mocap data integrates with your animation software

Motion capture software isn't always what your animators will work in. Many studios operate exclusively on Autodesk Maya, Blender, Unreal Engine, Unity, Cinema 4D Houdini, and others. Make sure the mocap data captured can integrate with your systems.

4. Capture motion and clean up the data

While motion capture data can be exceptionally accurate, it's not immune to errors. And that margin of error increases with erratic movements and high-speed motion. So be aware of the extra animation time you might need in post-production if your movements are complex. Don't forget to set aside a bit of time for you, and your animation team to clean up and refine the animation.

Advantages

Motion capture offers several advantages over traditional computer animation of a 3D model:

- Low latency, close to real time, results can be obtained. In entertainment applications this can reduce the costs of keyframe-based animation. The Hand Over technique is an example of this.
- The amount of work does not vary with the complexity or length of the performance to the same degree as when using traditional techniques. This allows many tests to be done with different styles or deliveries, giving a different personality only limited by the talent of the actor.
- Complex movement and realistic physical interactions such as secondary motions, weight and exchange of forces can be easily recreated in a physically accurate manner.
- The amount of animation data that can be produced within a given time is extremely large when compared to traditional animation techniques. This contributes to both cost-effectiveness and meeting production deadlines.
- Potential for free software and third-party solutions reducing its costs.

Disadvantages

- Specific hardware and special software programs are required to obtain and process the data.
- The cost of the software, equipment and personnel required can be prohibitive for small productions.

- The capture system may have specific requirements for the space in which it is operated, depending on camera field of view or magnetic distortion.
- When problems occur, it is easier to shoot the scene again rather than trying to manipulate the data. Only a few systems allow real-time viewing of the data to decide if the take needs to be redone.
- The initial results are limited to what can be performed within the capture volume without extra editing of the data.
- Movement that does not follow the laws of physics cannot be captured.
- Traditional animation techniques, such as added emphasis on anticipation and follow through, secondary motion or manipulating the shape of the character, as with squash and stretch animation techniques, must be added later.
- If the computer model has different proportions from the capture subject, artifacts may occur. For example, if a cartoon character has large, oversized hands, these may intersect the character's body if the human performer is not careful with their physical motion.

Applications

There are many applications of Motion Capture. The most common are for video games, movies, and movement capture, however there is a research application for this technology being used at Purdue University in robotics development.

Methods and systems

Reflective markers attached to skin to identify body landmarks and the 3D motion of body segments Silhouette tracking

Motion tracking or motion capture started as a photogrammetric analysis tool in biomechanics research in the 1970s and 1980s, and expanded into education, training, sports and recently computer animation for television, cinema, and video games as the technology matured. Since the 20th century, the performer has to wear markers near each joint to identify the motion by the positions or angles between the markers. Acoustic, inertial, LED, magnetic or reflective markers, or combinations of any of these, are tracked, optimally at least two times the frequency rate of the desired motion. The resolution of the system is important in both the spatial resolution and temporal resolution as motion blur causes almost the same problems as low resolution. Since the beginning of the 21st century - and because of the rapid growth of technology - new methods have been developed. Most modern systems can extract the silhouette of the performer from the background. Afterwards all joint angles are calculated by fitting in a mathematical model into the silhouette. For movements you can not see a change of the silhouette, there are hybrid systems available that can do both (marker and silhouette), but with less marker. In robotics, some motion capture systems are based on simultaneous localization and mapping.

Optical systems

Optical systems utilize data captured from image sensors to triangulate the 3D position of a subject between two or more cameras calibrated to provide overlapping projections. Data acquisition is traditionally implemented using special markers attached to an actor; however, more recent systems are able to generate accurate data by tracking surface features identified dynamically for each particular subject. Tracking a large number of performers or expanding the capture area is accomplished by the addition of more cameras. These systems produce data with three degrees of freedom for each marker, and rotational information must be inferred from the relative orientation of three or more markers; for instance shoulder, elbow and wrist markers providing the angle of the elbow. Newer hybrid systems are combining inertial sensors with optical sensors to reduce occlusion, increase the number of users and improve the ability to track without having to manually clean up data.

Passive markers

A dancer wearing a suit used in an optical motion capture systemMarkers are placed at specific points on an actor's face during facial optical motion capture.

Passive optical systems use markers coated with a retroreflective material to reflect light that is generated near the camera's lens. The camera's threshold can be adjusted so only the bright reflective markers will be sampled, ignoring skin and fabric.

The centroid of the marker is estimated as a position within the two-dimensional image that is captured. The grayscale value of each pixel can be used to provide sub-pixel accuracy by finding the centroid of the Gaussian.

An object with markers attached at known positions is used to calibrate the cameras and obtain their positions, and the lens distortion of each camera is measured. If two calibrated cameras see a marker, a three-dimensional fix can be obtained. Typically a system will consist of around 2 to 48 cameras. Systems of over three hundred cameras exist to try to reduce marker swap. Extra cameras are required for full coverage around the capture subject and multiple subjects.

Vendors have constraint software to reduce the problem of marker swapping since all passive markers appear identical. Unlike active marker systems and magnetic systems, passive systems do not require the user to wear wires or electronic equipment. Instead, hundreds of rubber balls are attached with reflective tape, which needs to be replaced periodically. The markers are usually attached directly to the skin (as in biomechanics), or they are velcroed to a performer wearing a full-body spandex/lycra suit designed specifically for motion capture. This type of system can capture large numbers of markers at frame rates usually around 120 to 160 fps although by lowering the resolution and tracking a smaller region of interest they can track as high as 10,000 fps.

Active marker

Body motion capture

Active optical systems triangulate positions by illuminating one LED at a time very quickly or multiple LEDs with software to identify them by their relative positions, somewhat akin to celestial navigation. Rather than reflecting light back that is generated externally, the markers themselves are powered to emit their own light. Since the inverse square law provides one quarter of the power at two times the distance, this can increase the distances and volume for capture. This also enables a high signal-to-noise ratio, resulting in very low marker jitter and a resulting high measurement resolution (often down to 0.1 mm within the calibrated volume).

The TV series *Stargate SG1* produced episodes using an active optical system for the VFX allowing the actor to walk around props that would make motion capture difficult for other non-active optical systems.

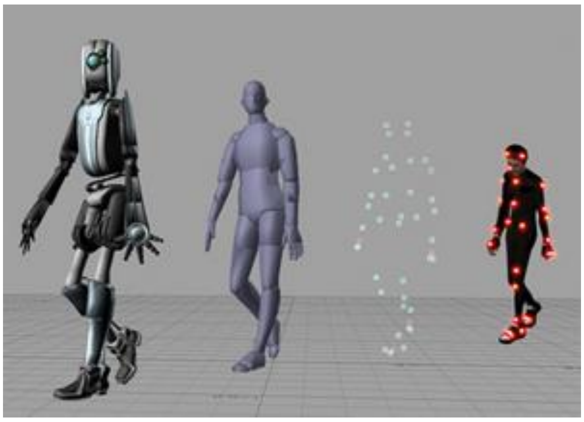
ILM used active markers in *Van Helsing* to allow capture of Dracula's flying brides on very large sets similar to Weta's use of active markers in *Rise of the Planet of the Apes*. The power to each marker can be provided sequentially in phase with the capture system providing a unique identification of each marker for a given capture frame at a cost to the resultant frame rate. The ability to identify each marker in this manner is useful in real-time applications. The alternative method of identifying markers is to do it algorithmically requiring extra processing of the data.

There are also possibilities to find the position by using colored LED markers. In these systems, each color is assigned to a specific point of the body.

One of the earliest active marker systems in the 1980s was a hybrid passive-active mocap system with rotating mirrors and colored glass reflective markers and which used masked linear array detectors.

A high-resolution uniquely identified active marker system with 3,600 × 3,600 resolution at 960 hertz providing real time submillimeter positions.

Time modulated active marker



Active marker systems can further be refined by strobing one marker on at a time, or tracking multiple markers over time and modulating the amplitude or pulse width to provide marker ID. 12-megapixel spatial resolution modulated systems show more subtle movements than 4-megapixel optical systems by having both higher spatial and temporal resolution. Directors can see the actor's performance in real-time, and watch the results on the motion capture-driven CG character. The unique marker IDs reduce the turnaround, by eliminating marker swapping and providing much cleaner data than other technologies. LEDs with onboard processing and radio synchronization allow motion capture outdoors in direct sunlight while capturing at 120 to 960 frames per second due to a high-speed electronic shutter. Computer processing of modulated IDs allows less hand cleanup or filtered results for lower operational costs. This higher accuracy and resolution requires more processing than passive technologies, but the additional processing is done at the camera to improve resolution via subpixel or centroid processing, providing both high resolution and high speed. These motion capture systems typically cost $20,000 for an eight-camera, 12-megapixel spatial resolution 120-hertz system with one actor.

IR sensors can compute their location when lit by mobile multi-LED emitters, e.g. in a moving car. With Id per marker, these sensor tags can be worn under clothing and tracked at 500 Hz in broad daylight.

Semi-passive imperceptible marker

One can reverse the traditional approach based on high-speed cameras. Systems such as Prakash use inexpensive multi-LED high-speed projectors. The specially built multi-LED IR projectors optically encode the space. Instead of retro-reflective or active light emitting diode (LED) markers, the system uses photosensitive marker tags to decode the optical signals. By attaching tags with photo sensors to scene points, the tags can compute not only their own locations of each point, but also their own orientation, incident illumination, and reflectance.

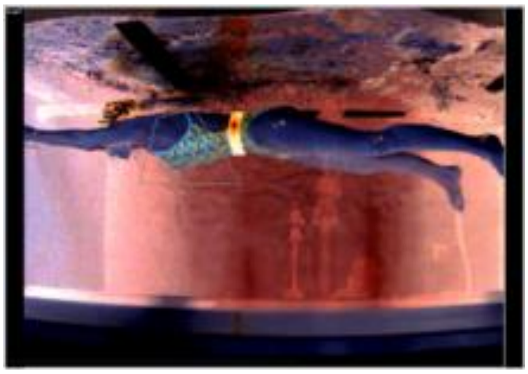
These tracking tags work in natural lighting conditions and can be imperceptibly embedded in attire or other objects. The system supports an unlimited number of tags in a scene, with each tag uniquely identified to eliminate marker reacquisition issues. Since the system eliminates a high-speed camera and the corresponding high-speed image stream, it requires significantly lower data bandwidth. The tags also provide incident illumination data which can be used to match scene lighting when inserting synthetic elements. The technique appears ideal for on-set motion capture or real-time broadcasting of virtual sets but has yet to be proven.

Underwater motion capture system

Motion capture technology has been available for researchers and scientists for a few decades, which has given new insight into many fields.

Underwater cameras

The vital part of the system, the underwater camera, has a waterproof housing. The housing has a finish that withstands corrosion and chlorine which makes it perfect for use in basins and swimming pools. There are two types of cameras. Industrial high-speed cameras can also be used as infrared cameras. Infrared underwater cameras come with a cyan light strobe instead of the typical IR light for minimum fall-off underwater and high-speed cameras with an LED light or with the option of using image processing.



Underwater motion capture cameraMotion tracking in swimming by using image processing

Measurement volume

An underwater camera is typically able to measure 15–20 meters depending on the water quality, the camera and the type of marker used. Unsurprisingly, the best range is achieved when the water is clear, and like always, the measurement volume is also dependent on the number of cameras. A range of underwater markers are available for different circumstances.

Tailored

Different pools require different mountings and fixtures. Therefore, all underwater motion capture systems are uniquely tailored to suit each specific pool instalment. For cameras placed in the center of the pool, specially designed tripods, using suction cups, are provided.

Markerless

Emerging techniques and research in computer vision are leading to the rapid development of the markerless approach to motion capture. Markerless systems such as those developed at Stanford University, the University of Maryland, MIT, and the Max Planck Institute, do not require subjects to wear special equipment for tracking. Special computer algorithms are designed to allow the system to analyze multiple streams of optical input and identify human forms, breaking them down into constituent parts for tracking. ESC entertainment, a subsidiary of Warner Brothers Pictures created especially to enable virtual cinematography, including photorealistic digital look-alikes for filming *The Matrix Reloaded* and *The Matrix Revolutions* movies, used a technique called Universal Capture that utilized 7 camera setup and the tracking the optical flow of all pixels over all the 2-D planes of the cameras for motion, gesture and facial expression capture leading to photorealistic results.

Traditional systems

Traditionally markerless optical motion tracking is used to keep track of various objects, including airplanes, launch vehicles, missiles and satellites. Many such optical motion tracking applications occur outdoors, requiring differing lens and camera configurations. High-resolution images of the target being tracked can thereby provide more information than just motion data. The image obtained from NASA's long-range tracking system on the space shuttle Challenger's fatal launch provided crucial evidence about the cause of the accident. Optical tracking systems are also used to identify known spacecraft and space debris despite the fact that it has a disadvantage compared to radar in that the objects must be reflecting or emitting sufficient light.

An optical tracking system typically consists of three subsystems: the optical imaging system, the mechanical tracking platform and the tracking computer.

The optical imaging system is responsible for converting the light from the target area into a digital image that the tracking computer can process. Depending on the design of the optical tracking system, the optical imaging system can vary from as simple as a standard digital camera to as specialized as an astronomical telescope on the top of a mountain. The specification of the optical imaging system determines the upper limit of the effective range of the tracking system.

The mechanical tracking platform holds the optical imaging system and is responsible for manipulating the optical imaging system in such a way that it always points to the target being tracked. The dynamics of the mechanical tracking platform combined with the optical imaging system determines the tracking system's ability to keep the lock on a target that changes speed rapidly.

The tracking computer is responsible for capturing the images from the optical imaging system, analyzing the image to extract the target position and controlling the mechanical tracking platform to follow the target. There are several challenges. First, the tracking computer has to be able to capture the image at a relatively high frame rate. This posts a requirement on the bandwidth of the image-capturing hardware. The second challenge is that the image processing software has to be able to extract the target image from its background and calculate its position. Several textbook image-processing algorithms are designed for this task. This problem can be simplified if the tracking system can expect certain characteristics that is common in all the targets it will track. The next problem down the line is controlling the tracking platform to follow the target. This is a typical control system design problem rather than a challenge, which involves modeling the system dynamics and designing controllers to control it. This will however become a challenge if the tracking platform the system has to work with is not designed for real-time.

The software that runs such systems is also customized for the corresponding hardware components. One example of such software is OpticTracker, which controls computerized telescopes to track moving objects at great distances, such as planes and satellites. Another option is the software SimiShape, which can also be used hybrid in combination with markers.

RGB-D cameras

RGB-D cameras such as Kinect capture both the color and depth images. By fusing the two images, 3D colored voxels can be captured, allowing motion capture of 3D human motion and human surface in real-time.

Because of the use of a single-view camera, motions captured are usually noisy. Machine learning techniques have been proposed to automatically reconstruct such noisy motions into higher quality ones, using methods such as lazy learning and Gaussian models. Such method generates accurate enough motion for serious applications like ergonomic assessment.

Non-optical systems

Inertial systems

Inertial motion capture technology is based on miniature inertial sensors, biomechanical models and sensor fusion algorithms. The motion data of the inertial sensors (inertial guidance system) is often transmitted wirelessly to a computer, where the motion is recorded or viewed. Most inertial systems use inertial measurement units (IMUs) containing a combination of gyroscope, magnetometer, and accelerometer, to measure rotational rates. These rotations are translated to a skeleton in the software. Much like optical markers, the more IMU sensors the more natural the data. No external cameras, emitters or markers are needed for relative motions, although they are required to give the absolute position of the user if desired. Inertial motion capture systems capture the full six degrees of freedom body motion of a human in real-time and can give limited direction information if they include a magnetic bearing sensor, although these are much lower resolution and susceptible to electromagnetic noise. Benefits of using Inertial systems include: capturing in a variety of environments including tight spaces, no solving, portability, and large capture areas. Disadvantages include lower positional accuracy and positional drift which can compound over time. These systems are similar to the Wii controllers but are more sensitive and have greater resolution and update rates.

They can accurately measure the direction to the ground to within a degree. The popularity of inertial systems is rising amongst game developers, mainly because of the quick and easy setup resulting in a fast pipeline. A range of suits are now available from various manufacturers and base prices range from $1000 to US$80,000.

Mechanical motion

Mechanical motion capture systems directly track body joint angles and are often referred to as exoskeleton motion capture systems, due to the way the sensors are attached to the body. A performer attaches the skeletal-like structure to their body and as they move so do the articulated mechanical parts, measuring the performer's relative motion. Mechanical motion capture systems are real-time, relatively low-cost, free from occlusion, and wireless (untethered) systems that have unlimited capture volume. Typically, they are rigid structures of jointed, straight metal or plastic rods linked together with potentiometers that articulate at the joints of the body. These suits tend to be in the $25,000 to $75,000 range plus an external absolute positioning system. Some suits provide limited force feedback or haptic input.

Magnetic systems

Magnetic systems calculate position and orientation by the relative magnetic flux of three orthogonal coils on both the transmitter and each receiver. The relative intensity of the voltage or current of the three coils allows these systems to calculate both range and orientation by meticulously mapping the tracking volume. The sensor output is 6DOF, which provides useful results obtained with two-thirds the number of markers required in optical systems; one on upper arm and one on lower arm for elbow position and angle. The markers are not occluded by nonmetallic objects but are susceptible to magnetic and electrical interference from metal objects in the environment, like rebar (steel reinforcing bars in concrete) or wiring, which affect the magnetic field, and electrical sources such as monitors, lights, cables and computers. The sensor response is nonlinear, especially toward edges of the capture area. The wiring from the sensors tends to preclude extreme performance movements. With magnetic systems, it is possible to monitor the results of a motion capture session in real time. The capture volumes for magnetic systems are dramatically smaller than they are for optical systems. With the magnetic systems, there is a distinction between alternating-current(AC) and direct-current(DC) systems: DC system uses square pulses, AC systems uses sine wave pulse.

Stretch sensors

Stretch sensors are flexible parallel plate capacitors that measure either stretch, bend, shear, or pressure and are typically produced from silicone. When the sensor stretches or squeezes its capacitance value changes. This data can be transmitted via Bluetooth or direct input and used to detect minute changes in body motion. Stretch sensors are unaffected by magnetic interference and are free from occlusion. The stretchable nature of the sensors also means they do not suffer from positional drift, which is common with inertial systems. Stretchable sensors, on the other hands, due to the material properties of their substrates and conducting materials, suffer from relatively low signal-to-noise ratio, requiring filtering or machine learning to make them usable for motion capture. These solutions result in higher latency when compared to alternative sensors.

Related techniques

Facial motion capture

Most traditional motion capture hardware vendors provide for some type of low-resolution facial capture utilizing anywhere from 32 to 300 markers with either an active or passive marker system.

All of these solutions are limited by the time it takes to apply the markers, calibrate the positions and process the data. Ultimately the technology also limits their resolution and raw output quality levels.

High-fidelity facial motion capture, also known as **performance capture**, is the next generation of fidelity and is utilized to record the more complex movements in a human face in order to capture higher degrees of emotion. Facial capture is currently arranging itself in several distinct camps, including traditional motion capture data, blend-shaped based solutions, capturing the actual topology of an actor's face, and proprietary systems.

The two main techniques are stationary systems with an array of cameras capturing the facial expressions from multiple angles and using software such as the stereo mesh solver from OpenCV to create a 3D surface mesh, or to use light arrays as well to calculate the surface normals from the variance in brightness as the light source, camera position or both are changed. These techniques tend to be only limited in feature resolution by the camera resolution, apparent object size and number of cameras. If the users face is 50 percent of the working area of the camera and a camera has megapixel resolution, then sub millimeter facial motions can be detected by comparing frames. Recent work is focusing on increasing the frame rates and doing optical flow to allow the motions to be retargeted to other computer generated faces, rather than just making a 3D Mesh of the actor and their expressions.

Radio frequency positioning

Radio frequency positioning systems are becoming more viable as higher frequency radio frequency devices allow greater precision than older technologies such as radar. The speed of light is 30 centimeters per nanosecond (billionth of a second), so a 10 gigahertz (billion cycles per second) radio frequency signal enables an accuracy of about 3 centimeters. By measuring amplitude to a quarter wavelength, it is possible to improve the resolution down to about 8 mm. To achieve the resolution of optical systems, frequencies of 50 gigahertz or higher are needed, which are almost as dependent on line of sight and as easy to block as optical systems. Multipath and reradiation of the signal are likely to cause additional problems, but these technologies will be ideal for tracking larger volumes with reasonable accuracy, since the required resolution at 100 meter distances is not likely to be as high. Many scientists believe that radio frequency will never produce the accuracy required for motion capture.

Researchers at Massachusetts Institute of Technology researchers said in 2015 that they had made a system that tracks motion by radio frequency signals.

Non-traditional systems

An alternative approach was developed where the actor is given an unlimited walking area through the use of a rotating sphere, similar to a hamster ball, which contains internal sensors recording the angular movements, removing the need for external cameras and other equipment. Even though this technology could potentially lead to much lower costs for motion capture, the basic sphere is only capable of recording a single continuous direction. Additional sensors worn on the person would be needed to record anything more.

Another alternative is using a 6DOF (Degrees of freedom) motion platform with an integrated omni-directional treadmill with high resolution optical motion capture to achieve the same effect. The captured person can walk in an unlimited area, negotiating different uneven terrains. Applications include medical rehabilitation for balance training, bio-mechanical research and virtual reality.

3D pose estimation

In 3D pose estimation, an actor's pose can be reconstructed from an image or depth map.

Motion capture examples

Here's what motion capture looked like for Gollum when filming the original Lord of the Rings:

And later, for the prequel The Hobbit, Benedict Cumberbatch gives a captivating performance that makes Smaug the dragon truly terrifying:

Motion capture for the Incredible Hulk realistically translates the actor's performance - no matter whether he's hulked up or human:

A motion capture example from the 2016 game "Paragon." Notice how this motion is capture in real-time, but the movements are slightly jerky? A good motion capture actor will be comfortable creating sharp movements that are easier for animators to clean up.

For example, the God of War game animators trained with mocap actors rigorously before capturing any motion data. The resulting fighting moves were crisp and dramatic, winning over gamers with great gameplay.

The best motion capture systems are always evolving

Reference:

Motion capture from:

1. https://www.rokoko.com/insights/what-is-motion-capture-and-how-does-it-work-in-2022
2. https://en.wikipedia.org/wiki/Motion_capture