INTRODUCERE

**Motion capture** (sometimes referred as **mo-cap** or **mocap**, for short) is the process of recording the [movement](https://en.wikipedia.org/wiki/Motion_(physics)) of objects or people. It is used in [military](https://en.wikipedia.org/wiki/Military_science), [entertainment](https://en.wikipedia.org/wiki/Entertainment), [sports](https://en.wikipedia.org/wiki/Sports), medical applications, and for validation of computer vision[[3]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-3) and robotics.[[4]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-4) In [filmmaking](https://en.wikipedia.org/wiki/Filmmaking) and [video game development](https://en.wikipedia.org/wiki/Video_game_development), it refers to recording actions of [human actors](https://en.wikipedia.org/wiki/Motion_capture_acting), and using that information to animate [digital character](https://en.wikipedia.org/wiki/Digital_character) models in 2D or 3D [computer animation](https://en.wikipedia.org/wiki/Computer_animation).[[5]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-5)[[6]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-twsBackstage-6)[[7]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-twsGuardian-7) When it includes face and fingers or captures subtle expressions, it is often referred to as **performance capture**.[[8]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-twsWired-8) In many fields, motion capture is sometimes called **motion tracking**, but in filmmaking and games, motion tracking usually refers more to [**match moving**](https://en.wikipedia.org/wiki/Match_moving).

SLIDE 2 - IMAGINI

Cele doua imagini:

* Two repetitions of a walking sequence of an individual recorded using a motion-capture system
* Motion capture of two [pianists](https://en.wikipedia.org/wiki/Pianist)' right hands playing the same piece (slow motion)

AVANTAJE

Motion capture offers several advantages over traditional [computer animation](https://en.wikipedia.org/wiki/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](https://en.wikipedia.org/wiki/Animation).[[10]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-Xsens_MVN_Animate_-_Products-10) The [Hand Over](https://en.wikipedia.org/wiki/Hand_Over) technique is an example of this.

A **Hand-Over** is a term used in the [animation](https://en.wikipedia.org/wiki/Animation) industry to refer to the process of adding finger and hand [motion capture](https://en.wikipedia.org/wiki/Motion_capture) data to the pre-existing full-body motion capture data, using a hand motion capture device.

* 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.[[11]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-11)
* 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.[[12]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-12)
* Potential for free software and third party solutions reducing its costs.

DEZAVANTAJE

* 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 it is operated in, 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](https://en.wikipedia.org/wiki/Squash_and_stretch) animation techniques, must be added later. **Squash and stretch** is the phrase used to describe "by far the most important"[[1]](https://en.wikipedia.org/wiki/Squash_and_stretch#cite_note-The_Illusion_of_Life-1):47 of the [12 basic principles of animation](https://en.wikipedia.org/wiki/12_basic_principles_of_animation), described in the book [*The Illusion of Life*](https://en.wikipedia.org/wiki/The_Illusion_of_Life) by [Frank Thomas](https://en.wikipedia.org/wiki/Frank_Thomas_(animator)) and [Ollie Johnston](https://en.wikipedia.org/wiki/Ollie_Johnston).

The principle is based on observation that only [stiff](https://en.wikipedia.org/wiki/Stiffness) objects remain inert during motion,[[1]](https://en.wikipedia.org/wiki/Squash_and_stretch#cite_note-The_Illusion_of_Life-1):47 while objects that are not stiff, although retaining overall volume, tend to change shape in an extent that depends on [inertia](https://en.wikipedia.org/wiki/Inertia) and [elasticity](https://en.wikipedia.org/wiki/Elasticity_(physics)) of the different parts of the moving object. To illustrate the principle, a half-filled flour sack dropped on the floor, or stretched out by its corners, was shown to be retaining its overall volume as determined by the object's [Poisson's ratio](https://en.wikipedia.org/wiki/Poisson%27s_ratio).

* 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.

TIPURI DE TEHNOLOGII

# TODO

APLICATII

[Video games](https://en.wikipedia.org/wiki/Video_game) often use motion capture to animate athletes, martial artists, and other in-game characters.[[13]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-13)[[14]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-GPro82-14) This has been done since the [Sega Model 2](https://en.wikipedia.org/wiki/Sega_Model_2) [arcade game](https://en.wikipedia.org/wiki/Arcade_game) [*Virtua Fighter 2*](https://en.wikipedia.org/wiki/Virtua_Fighter_2) in [1994](https://en.wikipedia.org/wiki/1994_in_video_gaming).[[15]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-15) By mid-1995 the use of motion capture in video game development had become commonplace, and developer/publisher [Acclaim Entertainment](https://en.wikipedia.org/wiki/Acclaim_Entertainment) had gone so far as to have its own in-house motion capture studio built into its headquarters.[[14]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-GPro82-14) [Namco](https://en.wikipedia.org/wiki/Namco)'s 1995 arcade game [*Soul Edge*](https://en.wikipedia.org/wiki/Soul_Edge) used passive optical system markers for motion capture.

During [*Game Developers Conference*](https://en.wikipedia.org/wiki/Game_Developers_Conference) 2016 in San Francisco [*Epic Games*](https://en.wikipedia.org/wiki/Epic_Games) demonstrated full-body motion capture live in Unreal Engine. The whole scene, from the upcoming game [*Hellblade*](https://en.wikipedia.org/wiki/Hellblade:_Senua%27s_Sacrifice) about a woman warrior named Senua, was rendered in real-time. The keynote[[21]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-21) was a collaboration between [*Unreal Engine*](https://en.wikipedia.org/wiki/Unreal_Engine), [*Ninja Theory*](https://en.wikipedia.org/wiki/Ninja_Theory), [*3Lateral*](https://en.wikipedia.org/wiki/3Lateral), [*Cubic Motion*](https://en.wikipedia.org/w/index.php?title=Cubic_Motion&action=edit&redlink=1), [*IKinema*](https://en.wikipedia.org/w/index.php?title=IKinema&action=edit&redlink=1) and [*Xsens*](https://en.wikipedia.org/wiki/Xsens).

# TODO

PASSIVE MARKERS

*Passive optical* systems use markers coated with a [retroreflective](https://en.wikipedia.org/wiki/Retroreflective) material to reflect light that is generated near the cameras 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](https://en.wikipedia.org/wiki/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

ACTIVE MARKERS

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 inverse square law provides one quarter the power at two times the distance, this can increase the distances and volume for capture. This also enables 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*](https://en.wikipedia.org/wiki/Stargate_SG1) and the movies [*Van Helsing*](https://en.wikipedia.org/wiki/Van_Helsing_(film)) *and* [*Rise of the Planet of the Apes*](https://en.wikipedia.org/wiki/Rise_of_the_Planet_of_the_Apes)were produced using an active optical system.

TIME MODULATED ACTIVE MARKERS



MARKERLESS

Emerging techniques and research in [computer vision](https://en.wikipedia.org/wiki/Computer_vision) are leading to the rapid development of the markerless approach to motion capture. Markerless systems such as those developed at [Stanford University](https://en.wikipedia.org/wiki/Stanford_University), [MIT](https://en.wikipedia.org/wiki/MIT), and the [Max Planck Institute](https://en.wikipedia.org/wiki/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

NON-OPTICAL SYSTEMS

INERTIAL SYSTEMS

Inertial motion capture[[25]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-25) technology is based on miniature inertial sensors, biomechanical models and sensor fusion algorithms.[[26]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-26) The motion data of the inertial sensors ([inertial guidance system](https://en.wikipedia.org/wiki/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. The popularity of inertial systems is rising amongst game developers,[[10]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-Xsens_MVN_Animate_-_Products-10) mainly because of the quick and easy set up resulting in a fast pipeline. A range of suits are now available from various manufacturers and base prices range from $1,000 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

MAGNETIC SYSTEMS

Magnetic systems calculate position and orientation by the relative magnetic flux of three orthogonal coils on both the transmitter and each receiver.[[27]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-NGen10Mag-27) 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. With magnetic systems, it is possible to monitor the results of a motion capture session in real time.[[27]](https://en.wikipedia.org/wiki/Motion_capture#cite_note-NGen10Mag-27) The capture volumes for magnetic systems are dramatically smaller than they are for optical systems

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.

The two main techniques are stationary systems with an array of cameras capturing the facial expressions from multiple angles and using software (such as 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.

RF POSITIONING

RF (radio frequency) positioning systems are becoming more viableas higher frequency RF devices allow greater precision than older RF technologies such as traditional [radar](https://en.wikipedia.org/wiki/Radar). The speed of light is 30 centimeters per nanosecond (billionth of a second), so a 10 gigahertz (billion cycles per second) RF 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