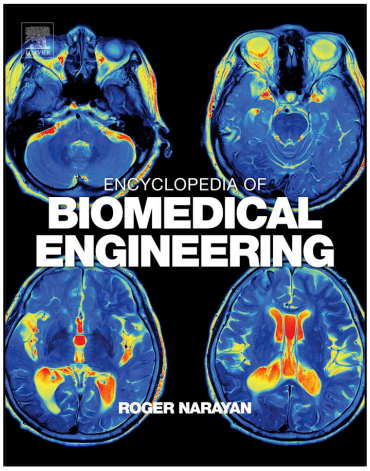
Author's personal copy

Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

This article was originally published in Encyclopedia of Biomedical Engineering, published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:
https://www.elsevier.com/about/our-business/policies/copyright/permissions

From Razavian, R. S., Greenberg, S., & McPhee, J. (2019). Biomechanics Imaging and Analysis. In R. Narayan (Ed.), Encyclopedia of Biomedical Engineering, vol. 2, pp. 488–500. Elsevier.
ISBN: 9780128048290
Copyright © 2019 Elsevier Inc. All rights reserved.
Elsevier

Author's personal copy

MEDICAL IMAGING

Biomechanics Imaging and Analysis

Reza Sharif Razavian, Sara Greenberg, and John McPhee, University of Waterloo, Waterloo, ON, Canada

© 2019 Elsevier Inc. All rights reserved.

Introduction: Motion Capture for Biomechanics

The science of measurement and analysis of the movement of humans and other animals constitutes a broad area in the field of biomechanics, which relies heavily on many imaging techniques. This article introduces the imaging tools that biomechanists use to record, analyze, and interpret the movement of humans. Specifically, two popular types of image-based motion capture systems are introduced: marker-based and markerless motion capture systems. This article also introduces the common workflow used by biomechanists for the processing of the motion capture data, in order to convert it to easily interpretable biomechanical variables.

Encyclopedia of Biomedical Engineering, 2019, 488–500

Author's personal copy

Types of Motion Capture Systems

Different position sensing technologies rely on certain physical measurements. Some are more direct (e.g., distance, velocity, or acceleration measurement), and some require further processing of the data (e.g., processing an image of an object). Image-based motion capture systems (also known as optical motion capture) are more popular for human motion recording in biomechanics compared with other methods such as accelerometers.

There are two general types of image-based techniques: (1) marker-based, in which 3-D position of special markers are calculated, and (2) markerless, in which certain features of an object within an image are identified and processed to obtain the object's 3-D position. Two or more cameras are usually involved to capture images of the objects from multiple angles, from which their 3-D position can be calculated.

The output of most motion capture systems (marker-based and markerless systems) is a set of 3-D points in space. They are either the marker positions that are placed on anatomical landmarks of the human body or the inferred positions of joint centers from markerless systems. For biomechanical applications such as inverse dynamics analysis, the motion of body segments (e.g., the hand, forearm, and upper arm) has to be calculated from the 3-D positions of markers or joint centers.

Camera Image Processing and Calibration

In a multicamera system, the 3-D position is calculated by processing multiple images of the same object taken from various angles. Usually, inclusion of more cameras enhances the accuracy and reliability of the measurement.

A point of interest on a camera sensor corresponds to a line of sight (Fig. 1); the physical position of the point is somewhere along this line. However, the distance from the camera cannot be calculated. A second image from another camera is needed to calculate depth. The intersection of the two lines of sights from the two cameras gives the location of the point in 3-D space.

Using homogeneous coordinates (instead of Cartesian coordinates) simplifies the calculations. The homogenous coordinates of a point in 3-D space are constructed by adding a fourth element to the triplet $[X, Y, Z]$ as $\mathbf{X} = [X, Y, Z, 1]$ (see Fig. 2). Likewise, the homogeneous coordinate of the image on the camera sensor plane is constructed as $\mathbf{u} = [u, v, 1]$. The transformation from the spatial position of an object to its image on the camera sensor in the homogeneous coordinates is given as

$$\mathbf{u}_{3\times 1} = \mathbf{P}_{3\times 4}\mathbf{X}_{4\times 1}$$

The camera parameter matrix, $\mathbf{P}_{3\times 4}$, is defined by 11 parameters that describe the position/orientation of the camera in space (6 extrinsic parameters) and its internal structure (5 intrinsic parameters relating to the focal length and sensor shape, skew, and reference frame).

A computer algorithm needs the camera parameter matrices to process 3-D positions of the objects. Such information is usually obtained using a calibration process that involves moving an object with known physical attributes in the capture volume. The raw camera images of the calibration object(s) in multiple frames are used as the input to the calibration process.

Each manufacturer has its own proprietary calibration algorithms. The common basis involves obtaining a measurement data matrix $\mathbf{w}$ by moving a marker within the capture volume:

$$\mathbf{w}_{3m\times n} = \begin{bmatrix} \mathbf{u}_1^1 & \cdots & \mathbf{u}_n^1 \\ \vdots & & \vdots \\ \mathbf{u}_1^m & \cdots & \mathbf{u}_n^m \end{bmatrix}_{3m\times n}$$

The matrix $\mathbf{w}_{3m\times n}$ contains the registered coordinates of the marker on the m camera sensors in multiple (n) frames. $\mathbf{u}_i^j = [u, v, 1]^T$ is the coordinates of the marker image in the i^{th} frame on j^{th} camera sensor. Given enough frames, the measurement matrix can be broken into two matrices $\mathbf{P}$ and $\mathbf{X}$ as

$$\mathbf{w}_{3m\times n} = \mathbf{P}_{3m\times 4}\mathbf{X}_{4\times n} = \begin{bmatrix} \mathbf{P}^1 \\ \vdots \\ \mathbf{P}^m \end{bmatrix}_{3m\times 4} [\mathbf{X}_1 \ \cdots \ \mathbf{X}_n]_{4\times n}$$

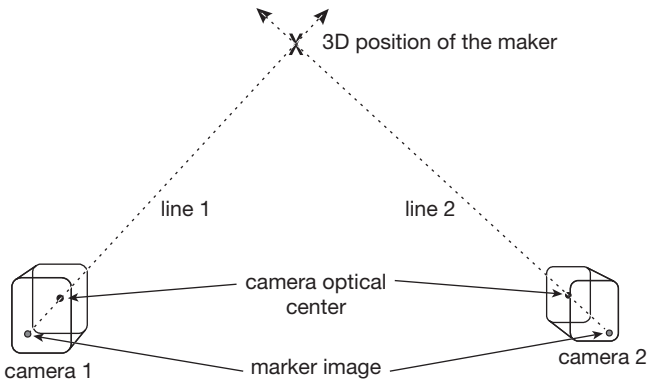


Fig. 1 Calculating 3-D position of a marker from its image on multiple cameras.

Author's personal copy

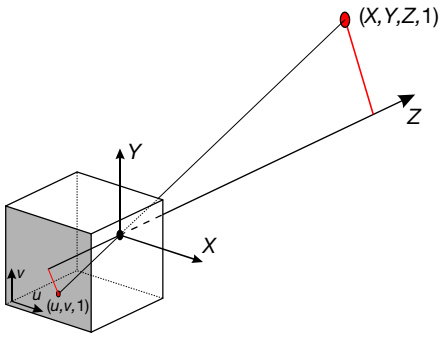


Fig. 2 A pinhole camera model and homogeneous coordinate systems.

where matrices $\mathbf{P}^j$ are the m camera parameter matrices and $\mathbf{X}_i$ are the inferred homogeneous 3-D coordinates of the marker in the corresponding frame. The calibration process may be improved using multiple markers with known relative distances, which facilitates the estimation of $\mathbf{P}$ and $\mathbf{X}$ matrices.

Marker-Based Motion Capture

By applying marker with known visual characteristics to a body, the body's motion can be inferred from the motion of the markers. These marker-based systems are known for their accuracy and robustness and, therefore, are the standard go-to solution when high position accuracy is needed. Besides the motion capture in biomedical research, animation and movie studios use marker-based motion capture systems extensively to reconstruct the same movements in virtual environments. Industries also use these techniques for measurements, although more accurate measurements such as from a coordinate measuring machine (CMM) are available.

Fig. 3 shows a typical setup, in which a number of cameras (firmly fixed) overlook a capture volume. Markers are usually small objects that are detected by the cameras. The image of the markers on the camera sensor plane is registered by the light-sensitive pixels. To improve robustness of measurements, infrared (IR) sensitive sensors are used instead of visible light sensors. The markers are either IR sources themselves (active markers) or reflective spherical objects that reflect the IR light emitted from an external source (passive markers).

The known coordinates of the pixels on the camera sensor plane are used for the processing. A pinhole camera model (Fig. 2) is usually used for the relation between an object's 3-D position and its image on the camera sensor. A 2-D Gaussian distribution can be used to obtain subpixel accuracy.

Active Versus Passive Markers

Active and passive markers are both used in commercial motion capture systems. Passive markers that reflect IR light from an external source have the advantage of being cheaper, lighter, and easier to install. Also there is no encumbrance due to wiring. However, the major disadvantage is the inability to explicitly distinguish markers from each other. In this setting, the markers are labeled manually in a starting frame, and a computer algorithm "follows" the markers in the subsequent frames by correlating their position with the previous frame. Therefore, marker mismatch is possible. Another disadvantage of passive marker systems is that artifacts may be registered by the cameras, because of other reflections in the capture volumes. Therefore, it is always recommended to cover reflective objects and surrounding walls to reduce these artifacts.

Active marker systems have the advantage of direct marker labeling. In active systems, a strobe mechanism turns the markers on and off at precise times, making the marker detection automatic. Furthermore, because the markers are light sources, the light can be of higher power compared with passive marker systems, resulting in higher range and accuracy. Lastly, registering artifacts (image of something other than the intended marker) is less likely in active marker systems. A disadvantage of the active marker systems is the heavier and more cumbersome markers. Line of sight is also an issue with active markers: a marker can be seen by a camera only if the camera is within the marker's angle of coverage, and the marker is within the camera field of view. This increases the possibility of marker loss, solely because the marker may not be facing the cameras. Lastly, the sampling rate is more limited with active markers, because all the markers must be turned on and off one by one in each frame (exemplified in Fig. 4). This results in a lower limit on the duration of each frame. Therefore, the maximum sampling frequency of active marker systems usually drops as more markers are used.

Standard Marker Sets for Biomechanics

There are recommended marker locations for biomechanical applications to ensure that enough information is captured to reconstruct the body motion. The International Society of Biomechanics (ISB) and some research institutes have published standard

Author's personal copy

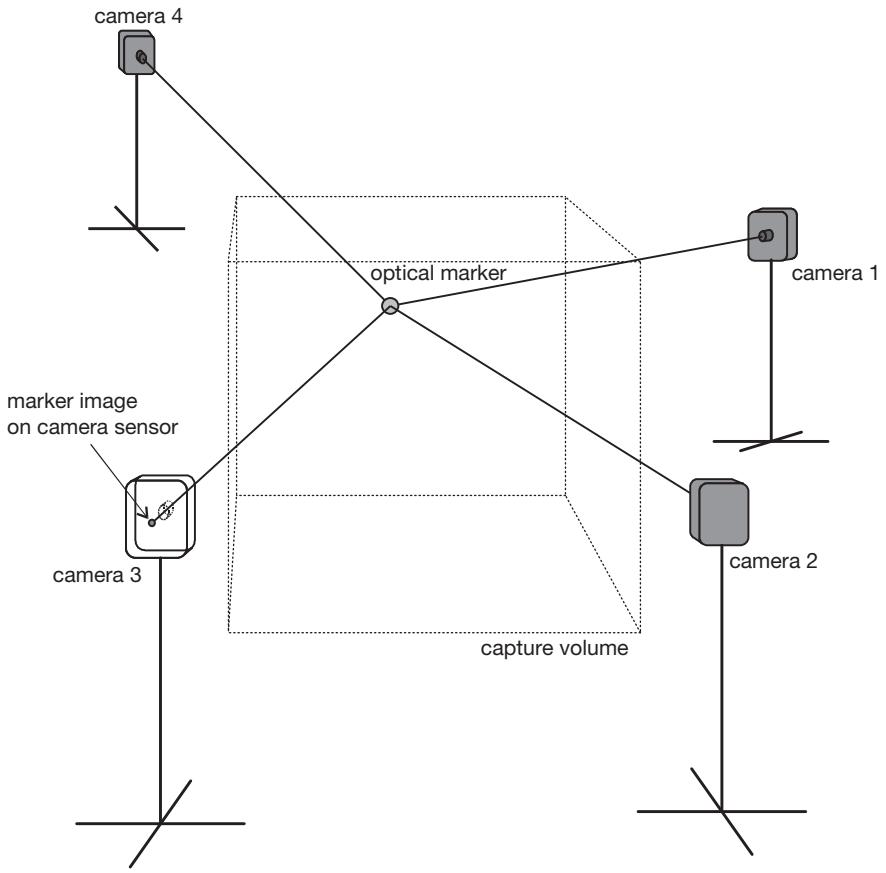


Fig. 3 A multicamera motion capture system.

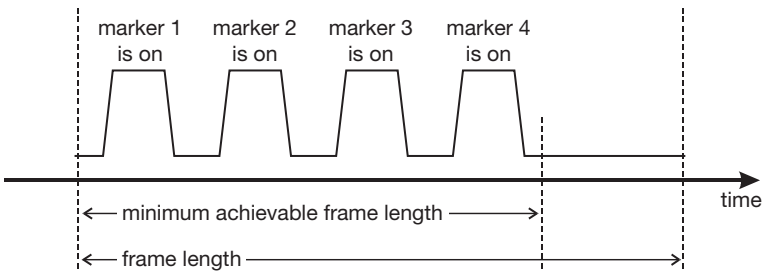


Fig. 4 Sampling rate is reduced with the number of markers used in active marker systems.

marker sets for various applications such as lower limbs, upper limbs, gait, and whole-body analysis. Table 1 shows a few examples of these marker sets. The table is not an exhaustive list; there are many more marker sets than are presented for various applications with different aims.

Marker Clusters

Placing markers on bony landmarks is a common way to capture limb motion. The ultimate goal of motion capture in biomechanics is often the calculation of joint rotations; therefore, the markers are distributed on a body segment as far apart as possible to increase accuracy. There are situations when such an approach is not feasible. For instance, to calculate femur rotation, the location of the hip joint center is required, which is not easy to estimate using surface marker locations. An alternative approach is to use a marker cluster that is affixed to the body segment. A marker cluster consists of a rigid base bearing a number of rigidly attached markers (at least three and noncollinear). Knowing the cluster dimensions, its location/orientation can easily and accurately be calculated. Attaching a marker cluster to a body segment can facilitate the calculation of its rotation.

Author's personal copy

Table 1 Examples of marker sets for different applications

Gait (Kadaba, 1990)	*Lower extremity (Wu et al., 2002)*	*Upper extremity (Wu et al., 2005)*	*Whole body (plug-in gait)*	
Sacrum (attached on a 10 cm wand)	PSIS[a]	Spinous process of the 7th cervical vertebra	Left/right front head (over the left/right temple)	Left/right ASIS
ASIS[b]	ASIS	Spinous process of the 8th thoracic vertebra	Left/right back head	Left/right PSIS
Greater trochanter	Lateral femoral epicondyle	Deepest point of incisura jugularis	7th cervical vertebrae	Left/right knee (lateral femoral epicondyle)
Lateral to knee joint	Medial femoral epicondyle	Most caudal point on the sternum	10th thoracic vertebrae	Left/right thigh (below the swinging of the hand)
Lateral malleolus	Medial malleolus	Most dorsal point on the acromioclavicular joint	Jugular notch	Left/right angle (lateral malleolus)
Between second and third metatarsal heads	lateral malleolus	Trigonum spinae scapulae	Xiphoid process of sternum	Left/right tibia (attached on a wand on the lower 1/3 of the shank)
Midthigh (attached on a 7 cm wand)	Most medial point of medial tibial condyle	Angulus inferior	Middle of the right scapula (no left marker)	Left/right toe (second metatarsal head)
Midshank (attached on 7 cm wand)	Most lateral point of lateral tibial condyle	Angulus acromialis	Left/right acromioclavicular joint	Left/right heel (calcaneus same height as toe)
	Tibial tuberosity	Most ventral point of processus coracoideus	Left/right upper arm between elbow and shoulder (placed asymmetrically)	
		Lateral humerus epicondyle	Left/right elbow (lateral humerus epicondyle)	
		Medial humerus epicondyle	Left/right forearm between the elbow and wrist (places asymmetrically)	
		Radial styloid process	Left/right wrist, thumb side (on a wand parallel to wrist axis)	
		Ulnar styloid process	Left/right wrist, pinkie side (on a wand parallel to wrist axis)	
			Left/right hand (second metacarpal head)	

[a]Posterior superior iliac spine.
[b]Anterior superior iliac spine.

In extreme situations when highest accuracy is needed, a marker cluster may directly be attached to the bone in a surgical operation. The cluster can therefore exactly represent the motion of the bone and remove the error due to skin movement and sensor placement.

Locating the Joint Centers

The locations of joint centers are important biomechanical variables that cannot be measured directly using motion capture systems (specifically, shoulder and hip joints are the hardest). To estimate their location, a number of methods are suggested.

Optimization-based methods

An approximation is to assume rigid body segments that are connected via ideal mechanical joints (e.g., the shoulder resembles a ball-and-socket joint). As a result, the distance of the markers from the joint center is assumed to remain constant. An optimization procedure can find the location of the joint center that minimizes the changes in the joint-to-marker distances.

Machine learning methods

The joints of the body are not ideal mechanical joints, and the distance of the markers from the joint centers is not necessarily constant and may be dependent on posture. To address this complex behavior, a model can be trained with machine learning methods. However, such a model requires training data, which is difficult to obtain.

Simple rules

It is also possible to estimate the joint center locations based on simple rules. For example, the midpoint between lateral and medial epicondyles of the humerus is often taken as the elbow joint center.

Author's personal copy

Sources of Error in Marker-Based Motion Capture and Mitigation Techniques

Soft tissue artifacts

Optical markers are usually attached to the skin using adhesive interfaces. As a result, there is a high chance of marker movement relative to the bone (due to skin stretch), deformation of the tissue underneath, and sensor/skin slippage. One study has reported that soft tissue artifacts may cause up to 16 mm relative movement between the marker and underlying bone.

Computer algorithms have been developed to partially compensate for marker movements. To achieve the highest accuracy, markers may be fixed directly to the bone via surgical operations to bypass soft tissue altogether. A noteworthy issue is the inertia of the sensor cluster, which may cause deflection in the attached pin and introduce error.

A noninvasive alternative is to use fluoroscopy to track moving bones directly. Fluoroscopy is a "video" of X-ray images; if taken from multiple angles, 3-D motion of bones can be calculated. The drawbacks are expensive equipment, limited capture volume, and the significant doses of radiation involved.

Marker occlusion

A marker has to be seen by at least two cameras to calculate its 3-D position. Therefore, it is possible to lose sight because of obstruction (by the body or obstacles), bad marker orientation (more prominent in active markers), or simply moving out of the capture volume. Including more cameras and properly placing them reduces the risk of losing markers.

In the case that a marker is lost, the missing trajectory has to be filled. Interpolation methods (linear or higher order) may be used to fill in the gap and satisfy continuity in the data. Other sophisticated methods such as Kalman filters may also be used to estimate the lost data.

Markerless Motion Capture

Another category of motion capture system involves recording a video of the moving object with one or more cameras and calculating its motion by detecting features in the frames. Unlike the marker-based system in which distinctive marker locations are measured, the markerless systems only rely on image processing techniques to get the location of the features in the frames. Therefore, their accuracy is limited by the performance of the detection algorithms, and their usefulness depends on properties of the scene and the object of interest. On the other hand, they have many advantages compared with the marker-based systems. Since they only need a video of the moving object, the sensing can be completely remote from the scene. An example is ball tracking in sporting events. Furthermore, the objects need little to no preparation, making it a faster method for data collection. The method is less intrusive than marker-based methods and does not impede motion. Implementations of this method employ regular video cameras or webcams that are often less expensive than the cameras used in marker-based motion capture systems. Lastly, a single-camera video can be used to obtain some information (such as 2-D position) about the motion, a feature that is not available in the marker-based systems.

RGB and Range (Depth) Images

Capturing a scene using a red-green-blue (RGB) sensor yields a set of frames of visible light that may be used to obtain information about motion in two dimensions within a plane parallel to the camera sensor. Measuring depth is not possible unless multiple synchronized videos from various angles are used. The processing of an object's location using multiple videos is very similar to the multicamera systems described previously for marker-based systems.

Alternatively, instead of multiple RGB images, a "range image" can be used to get 3-D information about the objects in a scene. Unlike a regular image in which each pixel contains light intensity information, a pixel of a range image contains depth information. An easy and inexpensive method for creating a range image involves an emitter that projects a light pattern onto the object and a camera that captures an image of the patterns from a different angle. The distorted pattern viewed by the camera can be used to calculate the depth of each pixel in the field of view. As an alternative to a pattern, a sensor may be able to detect the rebounded light and calculate depth based on the light time of flight. In both cases, the light may be visible or invisible. Invisible light is beneficial as it does not interfere with visible lights that may be captured by other RGB cameras. Infrared lasers and sensors are commonly used (e.g., in Microsoft Kinect or Leap Motion controller).

Markerless Feature Detection

The images of the markers in marker-based systems are easily identifiable in the camera sensors. In the absence of such identifiable markers in markerless systems, another layer of processing is needed to identify the objects of interest in each frame. The human eye is adept at object recognition; computers, on the other hand, require complex algorithms to perform such analyses.

There exist a variety of algorithms that can be used for this purpose. A common practice is to identify the human body in a frame and then fit a framework or "skeleton" to it (Fig. 5). The skeleton is in the shape of a stick figure, with multiple body segments connected to each other at joints (either real or imaginary). Some computer programs use a calibration process such as a T-pose posture in order to better fit a subject-specific skeleton to the human body.

Author's personal copy

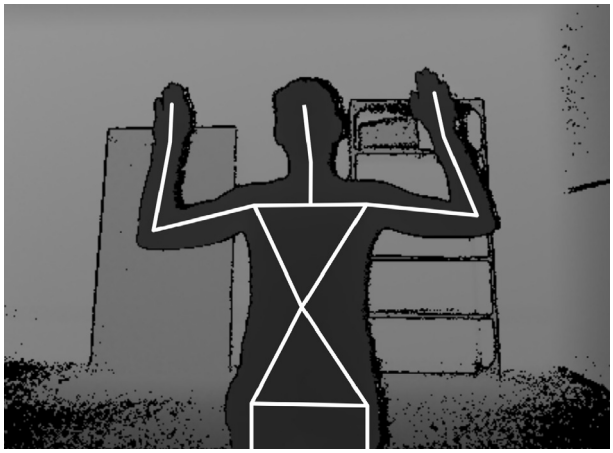


Fig. 5 The skeleton fitted to a body.

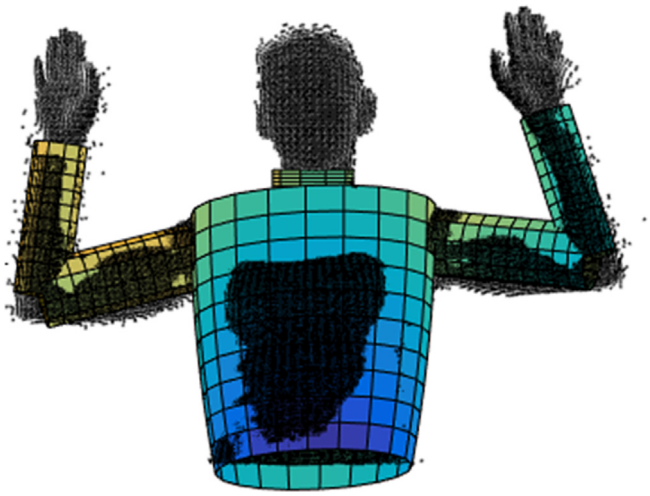


Fig. 6 Fitted 3-D shapes to the point cloud.

Fitting the modeled skeleton to the image can be accomplished by various image processing and machine learning techniques. Some methods involve matching 3-D shapes to the body segments (Fig. 6) or using learned models such as support vector machines, random forests, or artificial neural networks to identify joint locations.

Despite decreased accuracy, one advantage of markerless methods is that the body-segment rotations are readily available after the fitting process. In marker-based systems, the body-segment rotations must be calculated from skin markers, which are shown to be an error-prone process.

Image processing

There exists a multitude of image features that can aid in detecting a body. Often, these features are descriptors of points or sets of points within the image that may be robustly detected. Features may be detected in either an RGB image or a depth image.

One type of feature is an edge. An edge is a set of points at which there is a boundary between two segments of the image. This boundary is often described as a gradient of pixel values. A common tool for edge detection is the Sobel operator, which applies 3×3 convolutional filters to identify the magnitude of the image gradients in all directions at each pixel (Fig. 7). A Canny edge detector often uses the Sobel operator along with a Gaussian filter to reduce noise. A corner may be defined as the point of intersection of two edges or two gradients.

Another set of methods for identifying interest points in images are feature descriptors. These are numerical descriptions of shape, color, texture, and other identifiers that are able to be robustly detected. Feature detectors are used prominently in image stitching and object tracking in video and may make use of edges or corners. Feature or key points are detected in two or more frames, and these features are matched between frames based on their values and physical proximity within the image. Feature descriptors and detectors such as scale-invariant feature transform (SIFT) features, speeded-up robust features (SURF), and the histogram of oriented gradient (HOG) features are frequently implemented (Fig. 8).

Author's personal copy

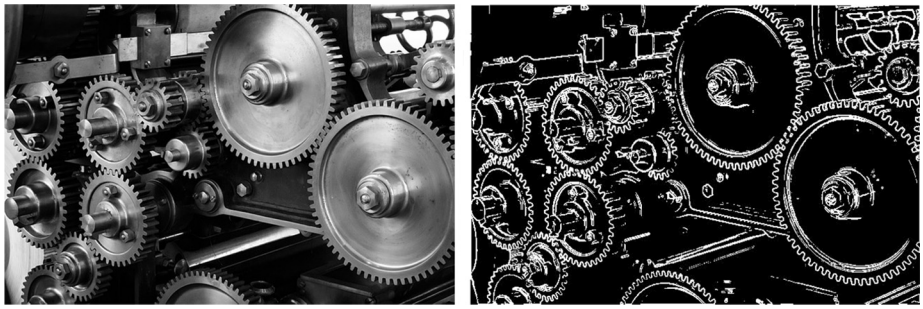


Fig. 7 A Sobel edge detector applied to an image of a series of gears.

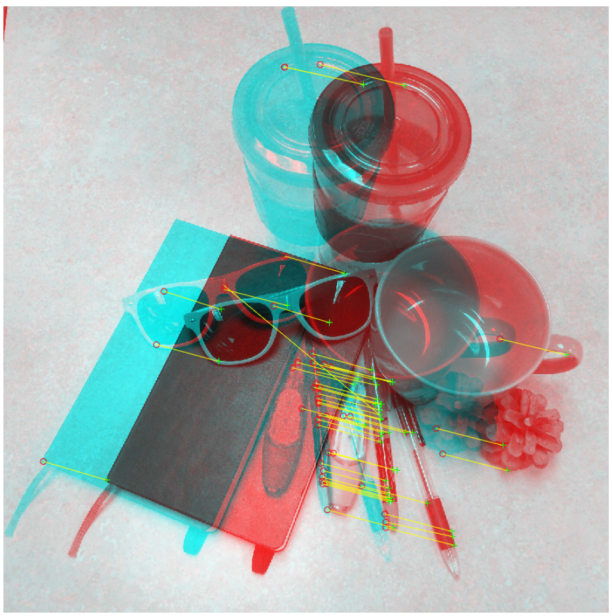


Fig. 8 Speeded-up robust features (SURF) matched between two frames with a translational difference.

Objects can also be identified through blob detection. To find a blob, the image is binarized according to one or more threshold values to identify only pixels having values within this new range (Fig. 9). If variance between the range of interest and the rest of the image is high, then automatic thresholding may be performed using Otsu's method of foreground segmentation. Once a binary image is acquired, the white areas representing the range of interest can be identified and classified according to their size and shape, allowing a specific blob to be identified if such information is known.

If a depth image or multiple RGB images are available, a point cloud of the scene may be captured. The methods discussed earlier may also be applied in three dimensions, or if the points belonging to the object or body of interest are known, then a model may be fitted to the body to identify its location and orientation. A common method of matching a model point cloud to the captured point cloud is the iterative closest point (ICP) algorithm, in which the points from each point cloud are matched to the closest correspondent, and translation and rotation are calculated to best align the model using these correspondences (see Fig. 6).

Machine learning

Machine learning techniques can also be applied to determine pose and gestures.

Support vector machines (SVMs) are learning models that are trained to classify data. Labeled data are used to construct a set of hyperplanes within the possible input space, which is then used to classify new data. The inputs can include interest points or features as described in the previous section.

Hidden Markov models (HMMs) are frequently implemented for gesture recognition. In a Markov chain, a system may exist in one of many states, with probabilities associated with changing from one state to another. In an HMM, many of these states are hidden from the observer. In the case of gesture recognition, it is impossible to manually identify the high-variance "states" of the body, so the states and probabilities are determined through training with a labeled data set of movements. After training, a particular gesture has a defined set of states, and the probability of a newly performed movement matching a gesture can be calculated using the HMM. A threshold probability decides whether or not a movement is classified as a defined gesture.

Author's personal copy

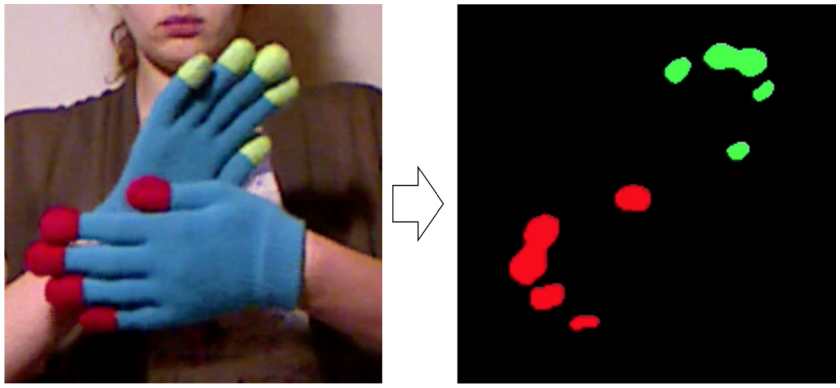


Fig. 9 An example of blob detection used to track fingertips.

The Microsoft Kinect, a combined RGB/depth camera, performs body pose recognition using randomized decision forests. This is an ensemble of decision trees trained independently, where each tree performs a classification or regression on a sample based on its features. The output of the randomized decision forest can be the mean or the mode of the tree outputs.

Artificial neural networks (ANNs) are also being applied to the problem of human body pose estimation and tracking. ANNs use an ensemble of digital units called neurons based on their biological namesake, arranged in layers, where the signals entering each unit are combined and adjusted according to a weight assigned during training. The output response is an estimation that can be used for a classification or regression. More complex versions such as convolutional neural networks have been applied successfully to two-dimensional pose estimation.

Sources of Error in Markerless Motion Capture and Mitigation Techniques

A big issue with markerless motion capture systems is the segmentation error. Correctly identifying each body segment is crucial to reconstructing the body motion. A loose shirt, for example, may cause serious challenges in correctly identifying various body parts.

Improving test conditions can greatly affect the results. Testing in a well-lit environment, high contrast between the subject and background (e.g., using different colors), and using tight clothing are a few remedies to improve the results.

More advanced machine learning techniques and inclusion of multiple cameras are shown to improve the outcome. However, it is still an ongoing research topic. There is a trade-off between robustness and performance for most methods, where the complexity of the model allows for increased precision and accuracy but is not always reliable or applicable to a large population.

Pros and Cons of Motion Capture Systems

Table 2 summarizes the common image-based motion capture systems for biomechanical applications and their advantages and disadvantages.

Other Imaging Techniques Used in Biomechanics

Fluoroscopy

Fluoroscopy is a method to track the motion of bones and other biological tissue in a series of X-ray images. Important drawbacks include expensive equipment, high doses of radiation, and small capture volume (even the state of the art is inapplicable for full-body motion tracking). Furthermore, a limited number of capture directions are available to reconstruct the 3-D motions (two directions are the state of the art). The nature of the radiations (X-ray) is also unfavorable, and long exposures are not possible.

Dual-Energy X-ray Absorptiometry

Another technology based on X-ray images is the dual-energy X-ray absorptiometry (DXA), which involves emitting two rays of X-ray with different energy content, in order to capture different tissues (bone and soft tissue mass, e.g., fat and muscle).

A DXA image has applications in biomechanics including calculation of bone and soft tissue mass and inertia properties of body segments.

Author's personal copy

Table 2 Comparison of motion capture systems

	Marker-based		*Markerless*		
	Active markers	*Passive markers*	*Single camera*	*Multiple camera*	*Depth camera*
Examples	Optotrak, Qualisys PhaseSpace	Vicon, OptiTrack Optotrak, Qualisys	Any camera	Organic motion, Simi, KinaTrax	Kinect, Leap
Pros	• Very high accuracy • Less postprocessing • No marker confusion	• Very high accuracy • Light markers • Unencumbering	• Easy to set up • Capture volume can be any size • Easy to calibrate (depends on application) • Portable • Inexpensive • Unencumbering	• Possibly high 3-D accuracy • Capture volume can be any size • Possibly inexpensive • Unencumbering	• Easy depth information • Usually comes precalibrated • Portable • Inexpensive • Unencumbering
Cons	• Larger markers • Sometimes wired markers restrict movements • Marker loss is more likely • Tedious calibration • Not very portable • Expensive	• Chance of marker confusion • Tedious calibration • Not very portable • Expensive	• No depth information • Requires image processing algorithm to obtain useful information	• Only as accurate as the image processing algorithm allows • Not very portable • Requires calibration	• Lower depth accuracy • Limited capture volume • Only depth from one angle

CT Scan

The last X-ray-based imaging technique used in biomechanics is a computed tomography (CT). A CT scan produces a 3-D image of the scanned object. One application of a CT image is the accurate 3-D reconstruction of bones to be used in computer programs such as finite element analysis.

MRI

Similar to a CT scan, magnetic resonance imaging (MRI) can be used to create 3-D images of the scanned object. Unlike a CT scan, an MRI machine uses very strong magnetic fields and nonionizing radio waves to change and measure the properties of a hydrogen nucleus; therefore, an MRI creates images of water content in living tissues.

MRI images are widely used in biomechanical research. Applications include measurement of muscle fiber length and orientations, tendon and ligament length and attachment points, joint structure and geometry, and bone structure.

An issue with an MRI image (also a CT scan) is its complexity. A trained eye is needed to distinguish between different tissues in an image. Manually labeling each area in an image is a tedious task. Sophisticated computer algorithms are being developed to make image segmentation automatic.

Ultrasound

Ultrasound imaging is an inexpensive and safe alternative to X-ray imaging to capture biological tissues. Recent advancements in ultrasound imaging techniques have allowed biomechanical researchers to capture images of muscle fibers, tendons, and joints to be used in biomechanical analysis.

Other Nonimaging Methods

Besides the image-based techniques mentioned earlier, there are other methods to obtain the 3-D position of objects.

Inertial measurement units (IMUs) have gained considerable attention in recent years, as inexpensive systems to measure motion. An IMU measures the linear accelerations (three-axis accelerometer) and rotational velocities (three-axis gyroscope), which can be numerically integrated to obtain 3-D position/orientation of an object. Since the IMUs directly measure linear accelerations and rotational velocities, their outputs are generally better than accelerations and velocities obtained by taking the time derivative of position data. However, a major problem with the current technology of IMUs is the drift in the position level (gradual divergence of the calculated position from the actual value), which is a result of the noise in the integrated signals. Sensor fusion (using multiple IMUs, global positioning system (GPS), height sensor, magnetometer, etc.) and model-based techniques (e.g., extended Kalman filter) are examples of the approaches taken to reduce the drift in IMUs.

Author's personal copy

Magnetic motion capture systems use a magnetic field transmitter/sensor to calculate the position and orientation of the attached object. Short-range interference with metallic (ferromagnetic) objects and bulky sensor/transmitter are some drawbacks of this technique. An advantage is that no direct line of sight between sensor and transmitter is needed.

Direct measurement methods such as electronic goniometers are also used in biomechanical applications to measure joint angles. Goniometers, however, have the disadvantage of being limited to measuring rotation about one or two axes (3-D rotations are not possible with one goniometer). The physical encumbrance is another disadvantage of these types of measurement.

Kinematic Analysis: Joint Coordinate System, Euler Angles, and Rotation Matrix

For many biomechanical applications (e.g., inverse dynamics or clinical assessments), experimental data for the location of joint centers and segment rotations are needed. Furthermore, reporting such information in clinically relevant terms facilitates the interpretation of the data.

A systematic way to obtain the joint rotations starts by fixing a coordinate system (CS) to each segment (called body-segment coordinate system). Relative rotation of the distal segment's CS with respect to the proximal ones defines the joint rotation.

Recommendations for Body-Segment Coordinate Systems

The ISB has proposed standards for defining the body-segment CS that results in clinically relevant joint rotations and promotes consistency among researchers. The suggestions are published in two articles:

- ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion – Part I: Ankle, hip, and spine, by Wu et al., (2002), *Journal of Biomechanics*
- ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion – Part II: shoulder, elbow, wrist, and hand, by Wu et al. (2005), *Journal of Biomechanics*

Rotation Matrix

Relative rotation of two adjacent CSs can be fully expressed by a "rotation matrix," $R_{3\times3}$. The rotation matrix can be constructed as

$$R = \begin{bmatrix} i.I & i.J & i.K \\ j.I & j.J & j.K \\ k.I & k.J & k.K \end{bmatrix}$$

where (I, J, K) are the unit vectors of the proximal segment's CS and (i, j, k) are those of the distal segment's CS. This rotation matrix describes how the proximal CS should rotate to obtain the distal one, that is,

$$CS_{distal} = [R]\, CS_{proximal}$$

A rotation matrix is the most robust method to express the relative rotation of two CSs. However, they are mathematical entities that are hard to interpret (e.g., in a clinical assessment report). That is the reason why joint coordinate systems (JCS) and Euler/Cardan angles are more popular in biomechanics.

Euler/Cardan Angles

A rotation matrix has nine elements; however, there are only three rotational degrees of freedom. Therefore, a rotation matrix contains redundant information. Euler angles express the transformation between two CSs using a triad of sequential rotations.

For instance, the body-fixed (ZXZ) sequence is shown in **Fig. 10** and described as follows: starting from the original CS (X, Y, Z), the first Euler angle (ϕ) specifies the rotation about the Z axis, which results in a new CS (X_2, Y_2, Z_2). Next, the second Euler angle (θ) specifies the rotation about the new X_2, resulting in a third CS (X_3, Y_3, Z_3). Finally, the third Euler angle (ψ) rotates the last CS about its Z_3, which gives the target CS (x, y, z). An Euler angle representation is sequence-dependent, meaning that another sequence (e.g., YZX) results in a different coordinate system.

A sequence with the same first and third axes (e.g., XYX or ZXZ) is usually called an Euler sequence, while a sequence involving all three is called a Cardan sequence (e.g., XYZ or ZXY).

The Euler angles may also be expressed as space-fixed, in which all the rotations occur about the axes of a CS that is fixed in space (may or may not be the initial CS).

Joint Coordinate System

An alternative way to express the rotation of two adjacent segments is the joint coordinate system (JCS) method. Similar to Euler angles, the relative rotation of two body-segment CSs is expressed by three rotation angles. The difference is that the rotations are

Author's personal copy

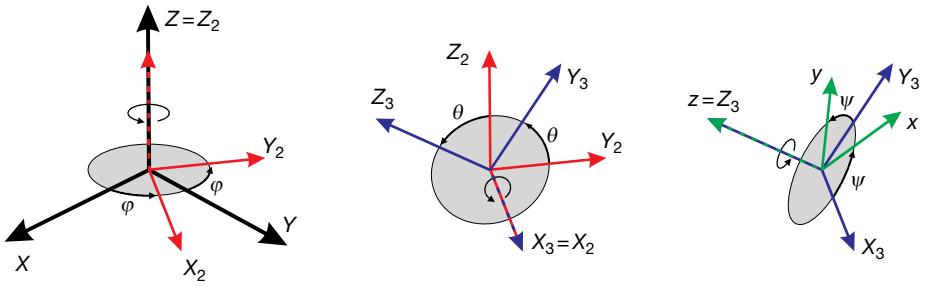


Fig. 10 A ZXZ Euler sequence.

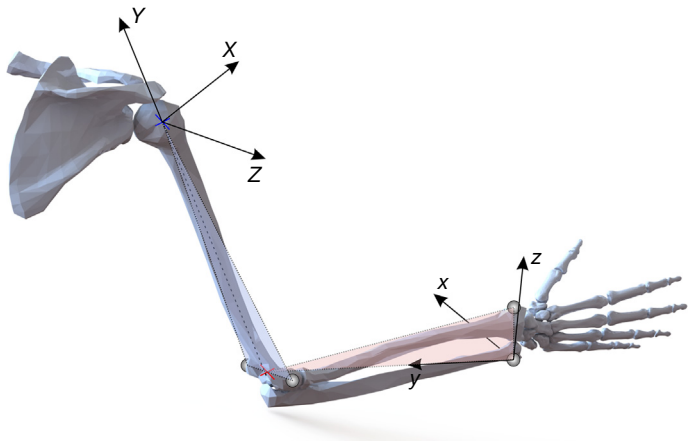


Fig. 11 The recommendation for arm segment coordinate systems.

about axes that belong to the distal and proximal CSs (as opposed to the sequential rotation about the axes of a single CS). Since the axes of rotation are taken from two separate CSs, their orthogonality is not guaranteed.

The JCS method is widely adopted, mostly due to its relevance to clinical terms. The example in the succeeding text is used to illustrate the JCS convention; it follows the ISB recommendations for JCS to obtain elbow rotations.

The first segment, humerus, has its CS (I, J, K) centered at the glenohumeral joint center (see Fig. 11). J is the unit vector of the line that connects the midpoint of medial and lateral epicondyles to the glenohumeral. I is perpendicular to the plane containing J and the two epicondyles. Finally, K is the cross product of $I \times J$.

The forearm CS (i, j, k) has its origin at the ulnar styloid process. The j vector is along the line connecting the origin to the midpoint of medial and lateral epicondyles. i is perpendicular to the plane that contains j, origin, and radial styloid process, pointing anteriorly (assuming body in standard anatomical position). The last unit vector is $k = i \times j$.

The elbow JCS is described as rotations about three axes (see Fig. 12); the recommendation for the first axis is $e_1 = K$ of the humerus CS; the second one is $e_3 = j$ of the forearm CS; the last axis, the "floating axis," is defined as the common perpendicular of the first two vectors $(e_f = e_3 \times e_1 = j \times K)$.

The first rotation of the elbow JCS (α about e_1) roughly gives the flexion angle and is calculated as the angle between the floating axis, e_f, and a reference axis fixed to the first body (taken to be I), that is, $\alpha = \cos^{-1}(e_f . I)$. Similarly, the rotation γ about e_3 is calculated as the angle between the floating axis and a reference axis on the second body (taken to be i). The angle $\gamma = \cos^{-1}(e_f . i)$, therefore, gives an approximation of the supination/pronation angle. The last rotation (abduction/adduction) is about the floating axis and is calculated as the angle between the two body-fixed axes $\beta = \frac{\pi}{2} - \cos^{-1}(e_1 . e_3)$.

Pros and Cons of Angle Calculation Methods

The Euler/Cardan angles and JCS are more or less similar and are both frequently used in biomechanical applications. It is often mentioned that the JCS is advantageous over Euler angles because JCS is sequence-independent. This statement may not be entirely accurate, as there is indeed a sequence that is embedded in the choice of the axes.

An important issue with the Euler/Cardan angles is the existence of a mathematical singularity (often called gimbal lock). This occurs when the second rotation is zero (Euler sequences) or 90 degrees (Cardan sequences). One can avoid singularities by using a quaternion (e.g., Euler parameters) instead of Euler angles, but quantities are difficult to interpret physically.

Author's personal copy

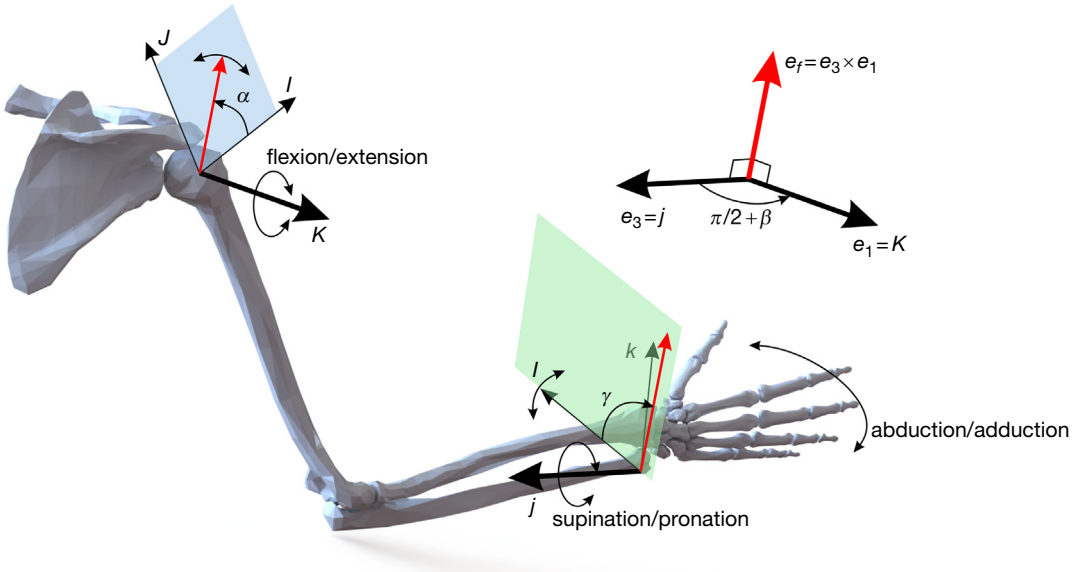


Fig. 12 Joint coordinate system to describe the elbow rotation.

Rotation matrices are the most complete form to express relative rotation of two segments but fail to provide tangible description about the rotations. It is hard to infer the amount of a joint rotation (e.g., knee flexion angle) directly from the elements of the rotation matrix without mathematical manipulations.

Conclusions

Image processing plays an important role in measuring human motions for biomechanics and clinical applications. Marker-based systems have been the dominant experimental modality to date, as a result of their superior positional accuracy. However, new modalities for human motion capture are being actively developed and promise lower cost and greater portability. In the future, we expect that markerless image processing will become the motion capture system of choice, once the issues of reliability and positional accuracy are resolved.

References

Kadaba, M. P., Ramakrishnan, H. K., & Wootten, M. E. (1990). Measurement of lower extremity kinematics during level walking. *Journal of Orthopaedic Research, 8*(3), 383–392. https://doi.org/10.1002/jor.1100080310.

Wu, G., Siegler, S., Allard, P., & Kirtley, C. (2002). ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—part I: ankle, hip, and spine. *Journal of Biomechanics, 35*(3), 474. https://doi.org/10.1006/rtph.2002.1549.

Wu, G., van der Helm, F. C. T., (DirkJan) Veeger, H. E. J., Makhsous, M., Van Roy, P., Anglin, C., & Buchholz, B. (2005). ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand. *Journal of Biomechanics, 38*(5), 981–992. https://doi.org/10.1016/j.jbiomech.2004.05.042.

Further Reading

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*, 679–698.

Alexander, E. J., & Andriacchi, T. P. (2001). Correcting for deformation in skin-based marker systems. *Journal of Biomechanics, 34*(3), 355–361. https://doi.org/10.1016/S0021-9290(00)00192-5.

Heikkila, J., & Silven, O. (1997). A four-step camera calibration procedure with implicit image correction. In *Computer vision and pattern recognition, 1997. Proceedings, 1997 IEEE. Computer Society Conference on* (pp. 1106–1112).

Shultz, R., Kedgley, A. E., & Jenkyn, T. R. (2011). Quantifying skin motion artifact error of the hindfoot and forefoot marker clusters with the optical tracking of a multi-segment foot model using single-plane fluoroscopy. *Gait & Posture, 34*(1), 44–48. https://doi.org/10.1016/j.gaitpost.2011.03.008.

Suntay, W. J. (1983). A joint coordinate system for the clinical description of three-dimensional motions: Application to the knee. *Journal of Biomechanical Engineering, 105*, 136.

Szeliski, R. (2010). *Computer vision: Algorithms and Applications*. Springer Science & Business Media.

Moeslund, T. B., Hilton, A., & Krüger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding, 104*(2–3), 90–126. https://doi.org/10.1016/j.cviu.2006.08.002.