Skeleton-Based Motion Capture for Robust Reconstruction of Human Motion

L. Herda, P. Fua, R. Plänkers, R. Boulic and D. Thalmann
Computer Graphics Lab (LIG)
EPFL
CH-1015 Lausanne, Switzerland
Email: {Lorna.Herda, Pascal.Fua, Ralf.Plaenkers, Ronan.Boulic, Daniel.Thalmann}@epfl.ch

Abstract

Optical motion capture provides an impressive ability to replicate gestures. However, even with a highly professional system there are many instances where crucial markers are occluded or when the algorithm confuses the trajectory of one marker with that of another. This requires much editing work on the part of the animator before the virtual characters are ready for their screen debuts. In this paper, we present an approach to increasing the robustness of a motion capture system by using a sophisticated anatomic human model. It includes a precise description of the skeleton's mobility and an approximated envelope. It allows us to accurately predict the 3–D location and visibility of markers, thus significantly increasing the robustness of the marker tracking and assignment, and drastically reducing— or even eliminating—the need for human intervention during the 3–D reconstruction process.

Keywords: Motion capture, skeleton-based tracking

1 Introduction

In recent years feature-length films have successfully exploited virtual actor technology. "Titanic" is one of the best known examples. It features hundreds of digital passengers with such level of realism that they are indistinguishable from real actors. The most critical element in the creation of digital humans was the replication of human motion: "No other aspect was as apt to make or break the illusion."[1] Optical motion capture offers a very attractive solution to this problem and provides an impressive ability to replicate gestures. Strolling adults, children at play and other lifelike activities have been recreated in this manner. The issues are slightly different for game-oriented motion capture. Capturing subtleties is less important because games focus more on big and broad movements. What matters more is the robustness of the reconstruction process and the amount of human intervention that is required.

In this last respect, the motion capture process is far from perfect. Even with a highly professional system there are many instances where crucial markers are occluded or when the algorithm confuses the trajectory of one marker with that of another. This requires much editing work on the part of the animator before the virtual characters are ready for their screen debuts.

In this paper, we present an approach to increasing the robustness of a motion capture system by using a sophisticated anatomic human model. It includes a precise description of the skeleton's mobility and an approximated envelope. It allows us to accurately predict the 3–D location and visibility of markers, thus significantly increasing the robustness of the marker tracking and assignment, and drastically reducing— or even eliminating—the need for human intervention during the 3–D reconstruction process. In contrast to commercially available approaches to motion capture such as the ones proposed by Elite^{tm} and VICON^{tm}, we do not treat 3–D marker reconstruction independantly from motion recovery. Instead we combine these two processes and use prediction techniques to resolve ambiguities. For example, we can predict whether or not a marker is expected to be occluded by the body in one or more images and take this knowledge into account for reconstruction purposes. When

1

a marker cannot be reconstructed with certainty from its image projections, we use the expected position of the skeleton to identify the marker and disambiguate its 3–D location. This is helpful when it is only seen by a small number of cameras. In our approach, the performer's skeleton motion is a byproduct of the reconstruction process.

In the remainder of this paper, we first review briefly some of the existing motion capture technologies. We then describe our skeleton-based approach. Finally we demonstrate its robustness using some complex motions that feature both large accelerations and severe occlusions.

2   Motion Capture Technologies

Motion capture technologies can be grouped into two broad classes.[7]

- **On-line motion capture.** The system's output can be directly used to pilot in real-time a virtual human body mimicking the performer's posture. This technology is often based on magnetic sensors.[2, 9] They are mainly used for Virtual Reality and on-line TV shows with synthetic characters.[7] However this technology is limited in several respects: Range of measurement space, noisy data, cumbersome sensors (although they tend to become smaller). As demonstrated at SIGGRAPH'99 by VICON^{tm} and Motion Analysis^{tm}, optical technology is becoming a serious contender in this area and can deliver results at a rate of 20 to 30 Hertz, provided that all markers remain visible.

- **Off-line motion capture.** Two processing stages are necessary to retrieve the performer's motion. This technology is typically based on optical motion capture from multiple camera views, usually in the infrared range. Despite the longer time required to visualize the captured motion, it is often preferred to magnetic technology. It allows the acquisition of the subtle gestures that are important in high-quality production to convey emotion through motion. It is used to capture the large and complex movements that are important in production to maintain a salient visual response to user input. It is also effective in a clinical context for the assessment of orthopaedic pathologies.

For both classes of technique, the character that is to be animated may differ considerably in shape and proportion from the performing artist. This can be solved using motion retargeting techniques.[6]

Optical technology is therefore bound to become increasingly useful for many applications. However traditional approaches such as the one depicted by Figure 1(a) suffer from occlusions and ambiguities that plague the 3–D reconstruction process. Typically, when markers become occluded, it grinds to a halt and requires user guidance. This limits its applicability in a real-time context and drives up post-processing costs for non real time applications. This is the issue that our proposed approach, depicted by Figure 1(b), addresses.

3   Approach

We use as input the 2–D camera data and calibration parameters provided by an Elite^{tm} optical motion capture system [5]. More precisely, as shown in Figure 2(b), we are given sets of 2–D point locations, one for each marker and each camera that sees it, and a projection matrix for each camera.

To extract a 3–D animation of a skeleton from a variety of movements performed by the same actor wearing the same markers, we first derive a skeleton-and-marker model, that is a skeleton scaled to the actor's body proportions and an estimate of the markers' locations with respect to the joints. To achieve this result, the actor is asked to perform a "Gym motion," that is a sequence of simple movements that involve all the major body joints. We can then use this calibrated skeleton for further motion capture sessions of more complex motions.

The complete approach is depicted by Figure 1(b).

3.1   Acquiring the Skeleton and Marker Model

During the calibration phase, our goal is to scale the bones of the generic skeleton of Figure 3(a) so that it conforms to the performer's anatomy and to model the marker's locations with respect to the joints. The complete skeleton, excluding detailed hands and feet, had 69 degrees of freedom (33 joints), plus six position parameters in 3–D space. The end result is a skeleton-and-marker model such as the one shown in Figure 3(b). In this work, we use a very simple marker
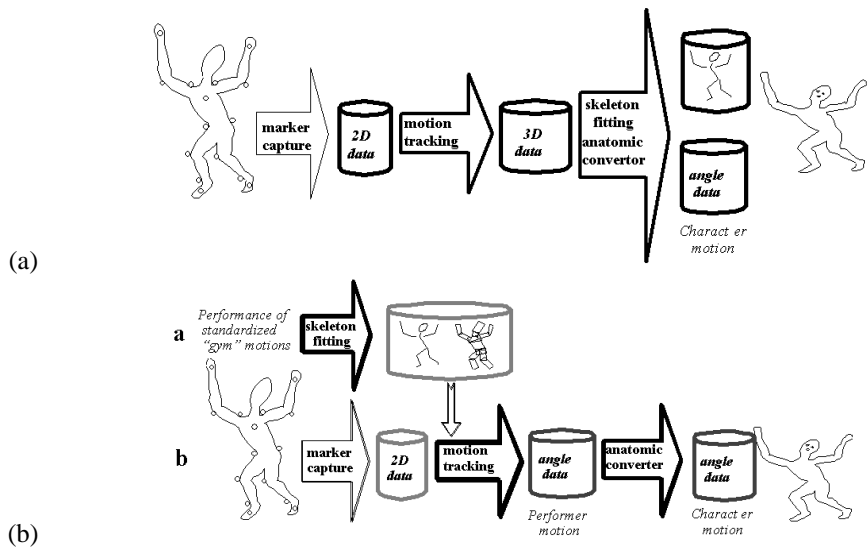
Figure 1. Approaches to Motion Capture: (a) In most commercially available packages, the estimation of the markers' 3–D positions and the fit of the 3–D skeleton are decoupled. (b) In our approach, we first compute a skeleton-and-marker model using a standardized set of motions. We then use it to resolve the ambiguities during the 3–D reconstruction process.
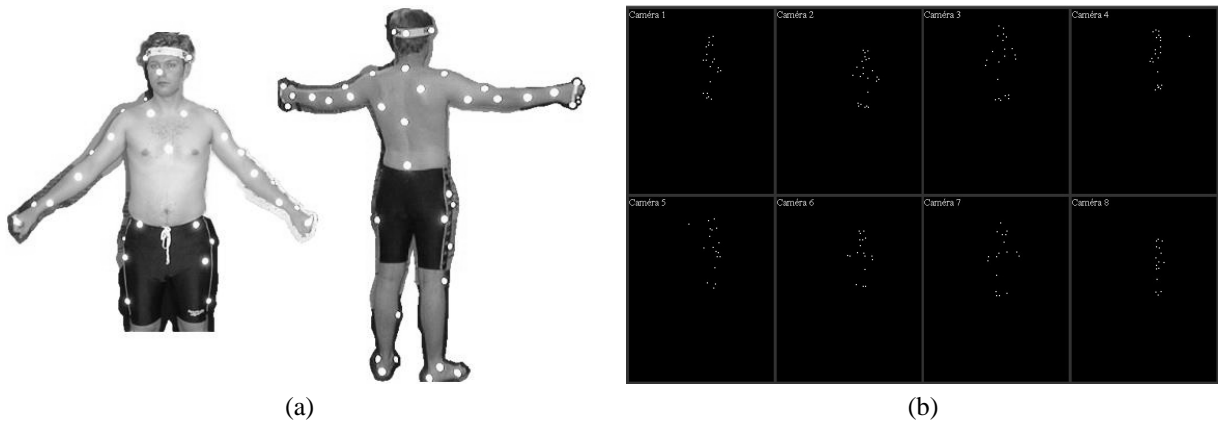

Figure 2. Input Data: (a) The performer wears markers and is imaged by eight infrared cameras. (b) For each camera, the Elite^{tm} system returns a 2–D location for each visible marker.

model: The markers are attached to specific joints and are constrained to remain on a sphere centered around that joint.

The skeleton-and-marker model is computed using least-squares minimization. As this is a non linear process, the system goes through three successive adjustment steps so as to move closer and closer to the solution at an acceptable cost while avoiding local minima. These steps are described below.

3.1.1 3–D marker reconstruction

As the gym motion is an especially simple routine highlighting the major joints motions, the 3–D location of the markers can be automatically and reliably reconstructed without knowledge of the skeleton for 200 to 300 frames at a time.

In practice, we partition the gym motion into independent sequences, each one involving only the mo-
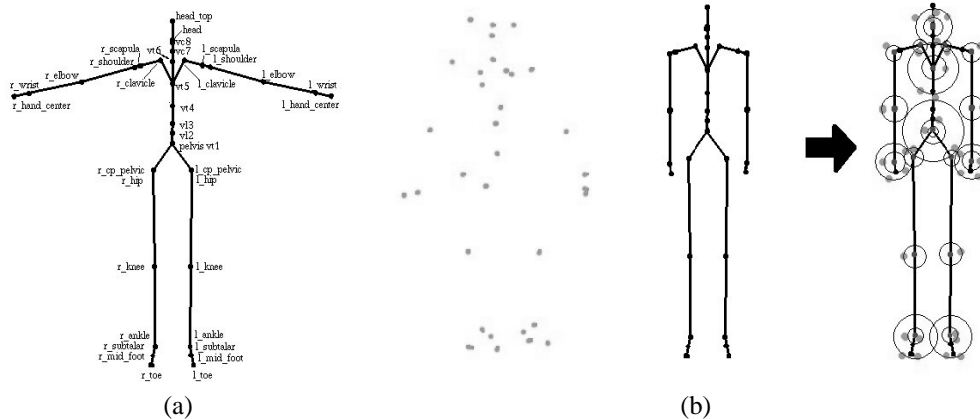
3

Figure 3. Skeleton and Marker Model. (a) Generic skeleton Model. (b) The generic model is scaled to conform the performer's anatomy. Each marker is attached to a joint and can move on a sphere centered around that joint.

tion of one limb or body part at a time. We then perform 3–D reconstruction and tracking for each one independently. If necessary, the user can reattach some markers to speficic body parts if they become lost.

3–D markers are reconstructed from the 2–D data using stereo triangulation [4]. In our examples, we use eight cameras. We first perform pairwise reconstruction. For each non-ambiguous stereo match, that is when there is only one possible candidate, we compute the corresponding 3–D coordinates on the basis of the 2–D coordinates. These 3–D coordinates are then re-projected onto the remaining six camera views, in order to determine the entire set of 2–D coordinates potentially associated with this one 3–D marker. We assume that a 3–D marker is correctly reconstructed if it re-projects into at least one other camera view, thus making a total of at least three camera views. We will say that these markers are reconstructed by trinocular stereo, that is, using at least three cameras. This is in contrast to markers reconstructed using only two camera views, and for which the projections into the other views failed.

Once we have reconstructed these trinocular 3–D markers in the first frame, we need to compare the number of reconstructed markers with the number of markers known to be carried by the actor. As all remaining processing is automatic, it is absolutely essential that all markers be identified in the first frame. Any marker not present in the first frame is lost for the entire sequence. Therefore, if the number of reconstructed markers is insufficient, a second stereo matching is performed, this time also taking into account markers seen in only two views. As binocular stereo matching is bound to introduce errors, the user is then prompted to confirm whether or not these binocular reconstructions are correct.

As soon as all markers are found in the first frame, the user is asked to associate each marker to a joint. For each highlighted marker, the user must select a body part and corresponding joint. Any marker not associated to a body part is discarded during the fitting process. Once these associations have been manually created, we can proceed with 2–D and 3–D tracking of the markers over the entire sequence.

2–D tracking is carried out at the same time as 3–D tracking because 2–D sequences are bound to provide more continuity than reconstructed 3–D sequences. We therefore use 2–D tracking in order to accelerate 3–D reconstruction: For each reliably reconstructed marker in frame [f], we consider the two sets of 2–D coordinates that were used to compute its 3–D coordinates. After 2–D tracking, these two sets of 2–D coordinates will most likely have links to two sets of 2–D coordinates in [f+1], the next frame. If so, we can then use them in [f+1] to construct the corresponding 3–D marker. To determine the related 2–D positions in the other camera views, we reproject the 3–D coordinates, as in the stereo matching process described above.

3–D tracking propagates the information attached to each marker in the first frame throughout the entire gym motion, so that as many markers as possible are

identified in all frames. A broken link in the tracked trajectory of a marker implies the loss of its identity and the user must then be prompted. In Section 3.2, we will see how we use the skeleton to overcome that problem in an automated fashion.

To compute the trajectory of a marker from frame [f] into frame [f+1], both in 2–D and 3–D, we look at the displacement of the marker over a four-frame sliding window [8]. The basic assumption is that displacement is minimal from one frame into the next, and the idea is to predict and confirm the position of a marker in the next frame. The displacement of a marker from [f-1] into [f] predicts the position in [f+1]. The actual position in [f+1] and the projection of the movement into [f+2] should confirm the previously-made hypothesis by eliminating ambiguities.

At the end of the marker reconstruction process and 2–D/3–D tracking steps, we have the gym motion reconstructed in 3–D, the trajectories of the markers throughout the sequence, as well as the identification of the markers with respect to the skeleton model.

3.1.2 Initial Joint Localization

In earlier work [11], we have developed a non-iterative technique that allows us to use these tracked markers to roughly estimate the 3–D location of a few key joints in each frame of the sequence, as well as the relative 3–D trajectories of the markers with respect to the underlying joints. We introduce this technique briefly below and refer the interested reader to our earlier publication for additional details.

Let us consider a referential bound to a bone represented as a segment. Under the assumption that the distance between markers and joints remains constant, the markers that are attached on adjacent segments move on a sphere centered on the joint that links the two segments. The position of a segment in space is completely defined by three points. Thus, if we have a minimum of three markers on a segment, we can define the position and orientation of that segment in space. Afterwards, we compute the movement of the markers on adjacent segments in the referential established by these markers and we estimate their centers of rotation.

To take advantage of this observation, we partition the markers into sets that appear to move rigidly and estimate the 3–D location of the center of rotation between adjacent subsets, which corresponds to the joint location.

This yields the approximate 3–D location of thirteen major joints, namely the joints of the arms and legs, as well as the location of the pelvic joint, at the base of the spine.

3.1.3 Skeleton Initialization

Given these thirteen joint locations in all frames, we take the median distances between them to be estimates of the length of the performer's limbs. We then use anthropometric tables to infer the length of the other skeleton segments.

This gives us a skeleton model scaled to the size of the actor. This model, however, is a static one, that is it has the appropriate dimensions but does not yet capture the postures for the gym sequence or the relative position of markers and joints.

To estimate those distances, we first need to roughly position the skeleton in each frame by minimizing the distance of the thirteen key joints to the corresponding centers of rotation. This is done by minimizing an objective function that is the sum of square distances from the centers of rotation to the joint it is attached to.

Given the fact that we use a sampling rate of 100 Hertz and that the gym motion is slow, the displacement from one frame to another is very small. Fitting is performed one frame at a time, and the initial parameter values for frame [f] are the optimised parameters obtained from the fitting in the previous frame [f-1]. As we only have thirteen observations for each frame, we do not attempt to estimate all of the skeleton's degrees of freedom. Only ten joints (shoulders, elbows, hips, knees, pelvic joint and the fourth spine vertebra) are active while all the others remain frozen. This yields the postures of the skeleton in all frames of the gym motion. In other words, we now have values of the global positioning vectors and degrees of freedom in each frame, as well as a better approximation to the limb lengths of the skeleton.

3.1.4 Global Fitting

We now have a skeleton model that is scaled to the size of the performing actor, but we are still missing a complete marker model, that is one that specifies where the markers are positioned on the actor's body and their distance to the joints to which they are attached. This is computed by performing a second least-squares minimization where the actual 3–D

5

marker locations become the data to which we intend to fit the skeleton.

Markers are not located exactly on the joints and the marker-to-joint distances must be estimated. To this end, we superimpose the markers' 3–D coordinates with the previously computed skeleton postures. In each frame, we then compute the distance from the marker to the joint and we take the median value of these distances to be our initial estimate of the marker-to-joint distance. Taking the marker model to be the distance from marker to joint means that the marker is expected to always be located on a sphere centered at the joint.

We now have all the information required to fit the skeleton model to the observation data. The initial state is given by the previously obtained skeleton postures. As we need to check that all markers are present and identified before fitting, we do it one frame at a time. Technically, the fitting process is similar to the one we used to fit models to stereo video sequences. [10] The interested reader is referred to [3] for details on the algorithm.

For each frame and for each marker, once the fitting is complete, the distance between marker and joint is stored. At the end of the gym motion sequence, we have as many such distances per marker as there are frames. The median value of these distances is an improved approximation of the marker-to-joint distance and becomes the final marker model.

3.2 Capturing Complex Motions

The resulting skeleton-and-marker model can now be applied to motions that we actually wish to capture. The procedure is very similar to the one used in the global fitting step of the previous section. However, we are now dealing with potentially complex motions. Consequently, even though 2–D and 3–D tracking ensure the identification of a large number of markers from one frame to another, ambiguities, sudden acceleration or occlusions will often cause breaks in the tracking links or erroneous reconstructions. For this reason, it has proved to be necessary to increase our procedure's robustness by using the skeleton to drive the reconstruction process, as discussed below.

The user is once again required to identify the markers in the first frame. However, he will no longer be associating 3–D markers to joints, but directly to 3–D markers located on the body model as computed during the calibration phase.

3.2.1 Skeleton Based Tracking

In order to improve the results of stereo matching, we use the skeleton for applying a visibility and occlusion test to each pair of 2–D markers used to construct a 3–D marker, thus verifying the validity of the reconstruction.

Visibility Check A marker is expected to be visible in a given view if it is seen more or less face on as opposed to edge on, that is if the surface normal at the marker's location and the line of sight form an acute angle. Suppose that we have reconstructed a certain 3–D marker using the 2–D pair (marker i1, view j1) and (marker i2, view j2); we check that these two markers i1 and i2 are indeed visible in views j1 and j2 respectively. Still assuming that displacement is minimal from one frame to the next, we use the the skeleton's posture in the previous frame and calculate the normal at the 3–D marker's location with respect to its underlying body part segment. We draw the line joining the 3–D marker coordinates to the position in space of the camera and if the angle between the normal and the line is acute, then the marker is visible. If this test shows that we have used the wrong 2–D coordinates for reconstruction, we must select other candidate 2–D coordinates: As discussed in Section 3.1.1, each 3–D marker is associated to two sets of 2–D coordinates determined by stereo correspondence, which we then use for reconstructing the 3–D marker. To this 3–D marker, we then also associate the 2–D coordinates from the remaining camera views onto which the 3–D coordinates of the marker projected correctly. Given that the visibility test has detected an erroneous 3–D reconstruction, we choose one of the 2–D coordinates computed via 3–D to 2–D projection, and calculate new 3–D coordinates. We then perform a new visibility test, and if this fails, we repeat the entire procedure.

Occlusion check Once a 3–D marker has passed the visibility test, it needs to undergo the occlusion check: We want to ensure that the 3–D marker is not occluded from some camera views by another body part. To this end, we approximate body parts by solids, cylinders for limbs and a sphere for the head. In the case of limbs, the cylinder's axis is the corresponding bone and the radius is the average joint-to-marker distance of the markers associated to this body part. In the case of the sphere, the centre is the mid-point of the segment. For each 3–D marker, a line is traced from the

marker to the position of the camera, and tested for intersection with all body part solids. In case of intersection with a solid, the marker is most likely occluded from this camera view. Therefore, we conclude that we have used erroneous 2–D coordinates for reconstruction. As before, we choose other 2–D coordinates and repeat the process.

3.2.2 Marker inventory

When all the markers have been reconstructed and tested, we can proceed with tracking and fitting. More specifically, for each frame, we perform 3–D reconstruction, tracking from the previous frame into the present one, identification of all markers, and finally, fitting of the skeleton-and-marker model to the observations. In order for the fitting to work correctly, all markers must be present in every frame. To ensure this, we carry out a marker inventory after 3–D reconstruction and before fitting.

Say we have just performed 3–D reconstruction using the 2–D data of frame [f], and we have thus obtained a set of markers. We then proceed with the following checks:

1. If the number of markers reconstructed using trinocular stereo is smaller than the actual number of markers worn by the actor, we perform binocular reconstruction and add the newly calculated coordinates to the already existing list of markers.

2. We perform 3–D tracking from [f-1] into [f], thus identifying a certain number of markers in [f], i.e. attaching them to their legitimate joint.

3. If all markers are still not found, we attempt to identify the 3–D markers that are still anonymous. We find all the skeleton's joints that are missing one or more markers. Assuming that displacement is minimal from one frame to another, we retrieve the coordinates of these joints in the previous frame, and calculate the distance from these joints to each remaining unidentified 3–D marker; the distance closest to the marker-to-joint distance specified by the marker model yields an association of the 3–D marker to that joint.

4. If the distance from marker to joint is larger than the distance specified by the marker model, we "bind" the coordinates of the 3–D marker to the
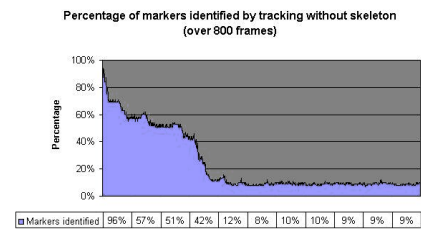

Figure 4. Percentage of markers identified by simple tracking, for the gym motion.
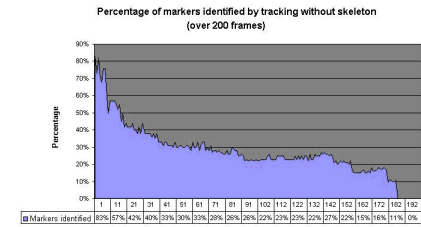

Figure 5. Percentage of markers identified by simple tracking, for a captured karate motion.

joint: We change its 3–D coordinates so that the marker moves within an acceptable distance of the joint. We however leave all reliably reconstructed 3–D markers untouched.

5. In the worst-case scenario, there may still be joints that are missing markers. We retrieve these markers in the three previous frame [f-3], [f-2] and [f-1], and calculate the acceleration; we apply this acceleration to the position in [f-1], thus obtaining an estimated position of the marker in the current frame [f]. As before, we calculate the distance from this inferred position to its associated joint. If it is out of range, we "bind" the coordinates.

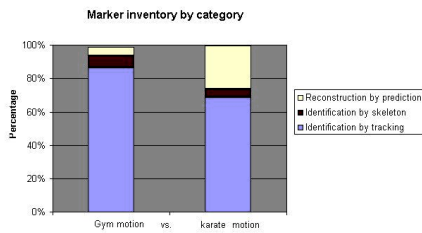In this manner, all 3–D markers are available for the fitting process.

7

Figure 6. Percentage of markers identified by each step of the marker inventory, respectively for the gym and the karate motion.

4 Results

4.1 Improvements provided by the skeleton-based tracking

In the following, we will give some statistics on the assistance provided by the skeleton with respect to tracking and marker reconstruction. In Fig. 4, we have run 3D tracking over 800 frames of the calibration gym motion. This tracking uses only simple marker prediction without skeleton information (see paragraph 3.1.1) over a sliding window of four frames. If a marker is lost in a frame, it cannot be recovered in any of the subsequent frames. The figure shows that after about 300 frames, the number of tracked markers drops to about 10% (tests performed with an actor using 32 body markers). In Fig. 5, we have performed the same statistics, but this time using a captured motion containing a fast movement, and we notice that simple tracking loses all markers in less than 200 frames.

With skeleton-based tracking, all markers are recovered in all frames, i.e. 100% of the markers are present and identified.

In Fig. 6, we show the percentage of markers identified by each process of the marker inventory of skeleton-based tracking(see paragraph 3.2.2). The two barcharts compare these percentages in the case of the gym motion and the karate motion. The first step of the process identifies reconstructed markers using simple 3D tracking. This percentage is obviously higher than in the previous figures, because in the skeleton-based tracking case, all markers are present in each frame, whereas in the simple tracking case, a marker lost in one frame is lost forever. The
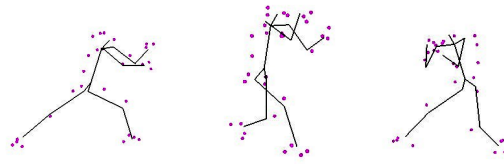

Figure 7. Three different frames from the karate motion (set 30 frames apart), seen from various viewpoints.
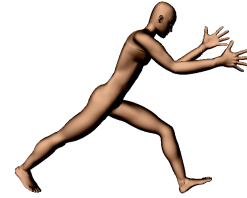

Figure 8. Virtual actor performing one of the recovered motions.

second step identifies reconstructed markers using the position of the skeleton in the previous frame. As to the third step, it reconstructs the markers that are still missing in the current frame, combining prediction of the 3D tracking type and the position of the underlying skeleton.

4.2 Fitting Results

The images in Fig. 7 show a few results obtained with one of the movements provided by the Motion Capture Studio. The use of the skeleton has enabled us to improve every step of the process, from 3D reconstruction, to tracking and identification of the markers. It is robust with respect to noisy data, out-of-bound and non-identified markers will be rejected, and it can also handle the case of occluded markers.

4.3 Discussion

Presently, motion capture software solutions mostly require human intervention for solving ambiguous stereo matches, as well as for re-identifying markers when a broken tracking link occurs. In our case, the entire process is automatic, as soon as initialization has been performed by the user.

As is the case of all methods, there is however space for improvement.

4.3.1 Pre-set thresholds and parameters

In the case of tracking and identification, one possibility would be to allow for dynamic search neighbourhoods which would render the algorithm more robust with respect to sudden accelerations in the movement. These neighbourhoods would be function of the acceleration, thus expanding their radius when the movement accelerates. The thresholds set for tracking and fitting are directly linked to the average displacement between two frames. Therefore, thresholds set at the beginning of the session in view of a regular movement at a certain speed are most often not optimal in case of a sudden acceleration. The markers will move out of the boundaries of the search neighbourhoods defined by these thresholds, this resulting either in rejection of these markers (in the tracking case) or in these markers' co-ordinates being modified so as to move them back closer to the skeleton joints (in the identification case). For example, a rapid movement involving an extension will result in an extension that is not as full as in the original motion.

4.3.2 User intervention

With respect to initialisation, one could do without user intervention if the actor were to adopt a specific pose at the beginning of the sequence this being the norm for calibration in the context of motion capture.

4.3.3 Gimbal lock

A problem we have encountered during the fitting phase is the well-known Gimbal lock problem that occurs due to the use of Euler angles for expressing rotations around the axes. The loss of one degree of rotational freedom results in the fact that a rotation doesn't occur due to the alignment of the axes. When this happens in our case, the output skeleton animation suffers sudden jumps from one position to another, when rotations do not occur over several frames. In the future, we could consider using quaternions to express the axis rotations.

4.3.4 Over-determination of the problem

Another problem regarding fitting is multiple solutions, namely in the case of the spine. Even when we keep only two joints in the spine, multiple solutions are possible for each posture, resulting in twisting at the spine level, and torsion or roll at the pelvic joint level. To solve this problem, we need to introduce some constraints on the spine on our side, or perhaps add more markers to the back of the actor. There are presently four markers on the torso: one at the bottom of the back, one at the level of the third vertebrae, one at the base of the neck, and one on the chest. This is obviously too few in order to determine a unique solution.

4.3.5 Marker model limitations

Regarding the marker model we are using for the moment, we could in the future use a more sophisticated model that would take into account the relative trajectories of the markers, calculated in the joint localization phase (see paragraph 3.1.2). The markers are presently free to evolve on a sphere centred in the joint, whereas in reality, their relative movement is much more limited. This would enable us to be more precise when it comes to identifying a marker for sure, and also for reconstructing a missing marker around a joint.

5 Conclusion

We have presented an approach to increasing the robustness of an optical motion tracking device. We use a body-model during the 3–D reconstruction and tracking process to assist the 3–D reconstruction of the markers, take visibility constraints into account and remove ambiguities. This greatly increases the motion capture sustem's robustness and decreases the need for human intervention.

The results shown in this paper were obtained using as input the data produced by a specific optical system. However, as we only use the markers' 2–D image locations and the camera calibration parameters, the approach is generic and could be incorporated in any similar system. This should lead to significant reductions in motion capture post-processing costs in a real production environment.

References

[1] Titanic Special Reprint. *Cinefex*, (72), December 1997.

[2] N. Badler, M. Hollick, and J. Granieri. Real-Time Control of a virtual Human Using Minimal Sensors. *Presence*, 2(1):1–5, 1993.

[3] N. D'Apuzzo, R. Plänkers, P. Fua, A. Gruen, and D. Thalmann. Modeling Human Bodies from Video Sequences. In *Electronic Imaging, SPIE Photonics West Symposium*, San Jose, CA, January 1999.

[4] O.D. Faugeras and L. Robert. What can two images tell us about a third one? *International Journal of Computer Vision*, (18):5–19, 1996.

[5] Ferrigno G. and Pedotti A. Elite: A digital dedicate hardware system for movement analysis via real-time tv signal processing. *IEEE Transactions on Biomedical Engineering*, BME-32(11), 1985.

[6] M. Gleicher. Retargeting Motion to New Characters. In *Computer Graphics, SIGGRAPH Proceedings*, pages 33–42, August 1998.

[7] R. Maiocchi. *Interactive Computer Animation*, chapter Character Animation Using Motion Capture, pages 10–39. Prentice Hall, 1996. N. Magnenat Thalmann and D. Thalmann, Eds.

[8] N. Malik, T. Dracos, and D. Papantoniou. Particle tracking in three-dimensional turbulent flows - Part II: Particle tracking. *Experiments in Fluids*, 15:279–294, 1993.

[9] T. Molet, R. Boulic, and D. Thalmann. Human Motion Capture Driven by Orientation Measurement. *Presence*, pages 77–83, June 1999.

[10] R. Plänkers, P. Fua, and N. D'Apuzzo. Automated Body Modeling from Video Sequences. In *ICCV Workshop on Modeling People*, Corfu, Greece, September 1999.

[11] M.-C. Silaghi, R. Plänkers, R. Boulic, P. Fua, and D. Thalmann. Local and global skeleton fitting techniques for optical motion capture. In *Workshop on Modelling and Motion Capture Techniques for Virtual Environments*, Geneva, Switzerland, November 1998.