

Motion capture and visualization of the hip joint with Dynamic MRI and optical systems

Nadia Magnenat-Thalmann

,MIRALab - University of Geneva

We present a methodology for motion tracking and visualization of the hip joint by combining MR images and optical motion capture systems. MRI is typically used to capture the subject's anatomy while optical systems are used to capture and analyse the relative movement between adjacent bones of the joint. Reflective markers are attached to the subject's skin and their trajectories are tracked and processed. However, the skin surface deforms while in motion due to muscle contraction leading to significant errors in the estimation of trajectories. In order to reduce these errors, we use MR images to capture both the anatomy and the trajectories of the bone. Prior to the scanning, we attach skin markers to the subject in order to analyse the markers displacements relative to the bone. We reconstruct the anatomical models of the subject and we compute the markers trajectories from the images. Using these calculated trajectories, we select the best markers configuration based on the criteria of markers displacements. The optimized configuration is used for recording external movements with the optical motion capture system. The resulting animation is mapped onto the virtual body of the subject including internal bones and the joint motion is visualized.

1. Introduction

We present a methodology to capture internal motions of the hip joint using optical motion capture and MR imaging. Optical motion capture systems are widely used in joint motion analysis by the biomechanics community, the movement of the markers is used to infer the underlying relative movements between two adjacent bones. The problem with these methods is that the skin surface moves over the underlying structures. The ability of MRI scanners to image the articulation dynamically and non-invasively in vivo opens the way to the efficient and accurate modeling and analysis of patient-specific joints. Therefore MRI is used to quantify the markers displacements and select the best markers configuration to be used with the optical system in order to reduce the errors of skin artifacts.

In our study, we focus on the hip articulation. The clinical application is to develop a pre-surgical planning tool for hip osteoarthritis (degenerative hip disease). Tissue-specific MRI protocols are set up and used to acquire volumes of healthy hips. Generic anatomical models of the joint are built interactively, validated by our medical partners and deformed automatically to match patient organs geometry.

On the other hand, dynamic MRI protocols are developed in order to capture the subject's internal motion. Reflective markers are attached to the subject's limb, the internal motion of the bone as well as the skin markers are tracked using a developed tracking algorithm. The skin markers trajectories are used to select the "best" markers configuration to be used with the optical system. Finally, the individualized models are animated to simulate realistic hip joint motions. The rest of this paper is organized as follows: First, we present our methodology for the anatomical modeling of the hip joint using static MRI, then; we introduce our algorithm for bone tracking using dynamic MRI, followed by the optimization of the skin markers configuration. The last section is dedicated to external motion capture and visualization of the joint.

2. Anatomical modeling of the hip joint

2.1. Image acquisition

Two healthy adult subjects (a female and a male) have undergone the MRI scanning. The acquisition was performed at HUG (Hôpitaux Universitaires de Genève) with a 1.5 T Intera station manufactured by Philips Medical systems. Four high-resolution MRI scans containing thin axial slices are obtained for each subject. The scanning ranges from the ilium crest to the knee based on an axial localizer. The scan is extended up to the knee in order to determine the anatomical axis of the femur to perform motions of the hip joint [KSM*02][MTKG02]. In the following, the acquisition protocol is detailed:

Coil : A four elements phased array coil used for body imaging to optimize Signal-to-Noise ratio.

Subject positioning : Supine, with legs immobilized to reduce leg movement.

Imaging protocol : Three series of T1-weighted spin echo images with TR/TE=578/18ms, Matrix/FOV=512x512/40cm and slice thickness=2mm to 10mm for bones and two series of selective water excitation gradient echo sequences with TR/TE =20/7ms, Matrix/FOV=256x256/20cm and slice thickness=1mm for cartilage.

2.2. Identification of anatomical structures

Segmentation is performed using a custom-written discrete snake procedure [KWT88] to extract the hip and femoral contours. On each MRI slice, an initial set of points is digitized along each articular curve with a coarse spacing of 1-2cm (Figure 1). The active contour is then used to best fit the actual boundary. This provides an accurate location of the bone contour sufficiently near the initialization curve.

Although the snakes have proven to achieve high accuracy while decreasing the time required for manual segmentation, manual corrections are necessary on the slices with fuzzy edges. Moreover, the segmentation is validated by the medical experts before the reconstruction process to ensure maximum precision in the 3D models.

2.3. Anatomical structures reconstruction

The Marching Cubes algorithm, originally proposed by W. Lorensen [LC87], is considered to be a standard approach to the problem of extracting iso-surfaces from a volumetric dataset. Many implementations are available both as part of commercial systems or as public domain software. We use the Visual Toolkit [vtk] implementation of the Marching Cubes algorithm to generate iso-surfaces from the segmented volume. The resultant polygonal surface is simplified with Schroeder decimation algorithm [SZL92]. This technique is based on multiple filtering passes that remove vertices passing a minimal distance or curvature angle by analyzing the geometry and topology of a triangle mesh locally. This decreases the total number of polygons while preserving intricate surface details.

The decimated polygonal surface is smoothed by adjusting the coordinates of the vertices using Laplacian smoothing. The effect is to "relax" the mesh, making the cells better shaped and the vertices more evenly distributed. The hip (femur) model is reconstructed with 42,944 vertices and 84,313 triangles, the femur model with 27,608 vertices and 53,534 triangles. Cartilage models are reconstructed with 5856 vertices and 11454 triangles for the acetabular cartilage and 10241 vertices and 20 360 triangles for the femoral cartilage (Figure 1. (c), (d))

2.4. Bone models individualization

Due to the large amount of textural information, noise, low contrast and resolution, bone segmentation is often a difficult task. Despite many researchers have tried to provide robust and fully automatic segmentation tools, no method has proven to be generic and manual corrections are generally required. Two main approaches for the segmentation are presented in the literature [Bro92]: the inter-patient registration, aiming at aligning patient images to segmented generic images, and the deformable models-based approach, aiming at matching a generic model to features in patient images. Inter-patient registration deals with similarity between images and, consequently, is sensible to noise and differences between the generic and the patient acquisition protocol. We use a method based on discrete deformable models to segment individual bones.

First, the generic model is elastically initialized with a landmark-based approach and Thin-Plate-Splines interpolation [RSS*96]. Then, the model is deformed automatically by optimization of an energy function which is composed of an external energy term, measuring the matching between the model and image edges, and an internal energy term that maintains a smooth and connected model. External energy is calculated from the MRI oriented gradient images and model normals. The internal energy derives from deformation spheres [LM01] that constraint model deformation. In order to avoid convergence into local minima, we use a multi-resolution approach and deformation spheres with decreasing radius (the model's deformability increases).

It takes on average one hour for a model with 40,000 vertices to deform (Pentium4, 2GHz). The automatic method has been successfully tested and validated on four different pelvis and femurs with the same parameters: deformability (radius from 2cm to 1cm), number of iterations (10000), number of landmarks (16) and number of resolutions (3). The difference between manual and automatic segmentation is less than 15% of the total number of voxels for bones (Figure 2)

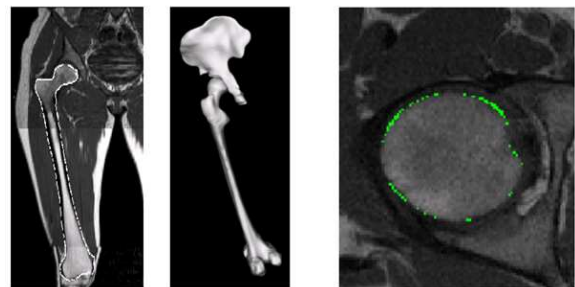


Figure 2: Automatic bone segmentation on a sample slice, corresponding 3D models and difference between manual and automatic segmentation on a sample slice (15%)

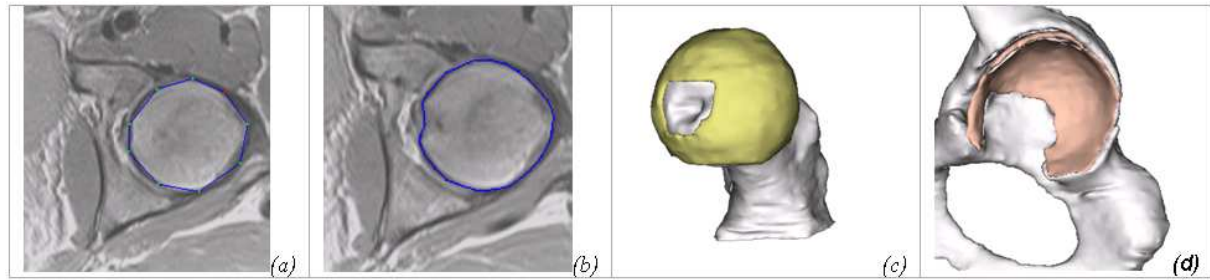


Figure 1: (a) Manual digitizing of the articular boundary (b) fitting of the active contour (c): Femoral head and femoral cartilage 3D surfaces. (d) acetabular cartilage and hip 3D surfaces.

3. Bone motion capture from MRI

3.1. Real-time dynamic images acquisition

The imaging protocol was developed and optimized with reference to the limitations of tracking algorithm. Firstly, the trade-off in image quality with FOV and matrix was investigated qualitatively on healthy volunteers in order to achieve the optimum resolution, contrast and frame acquisition time. As scan duration was proportional to the phase encode matrix, the phase encode matrix was maintained <100 at the shortest repetition time possible (TR 3.5ms). It was found that reducing the FOV and hence the phase encode matrix, maintaining an in-plane resolution of 2mm, was not an effective way to reduce frame acquisition time, due to the need to use fold-over suppression to avoid aliasing in the phase encode direction. A parallel imaging technique, SENSE (Philips Medical Systems, Best NL), was found to reduce the scan time by a factor of 2 without significant reduction in image quality. A reference scan is acquired prior to the SENSE MR sequence to measure the sensitivity profile of the phased-array coil. The same reference scan is used for all the images of the dynamic series. A positioning device was developed that facilitated reproducible abductive motion in both sequential and dynamic modes. A study was run with six healthy volunteers to optimize and evaluate the robustness of the registration-MRI protocol combination without the introduction of motion artifacts. Ethics approval was obtained from the local ethics committee for the study protocol. In a first session a complete static image data set of the pelvis and femur was acquired with a 2D multi-slice spin echo acquisition (TR/TE 578/18ms). In the second scan session the joint was stepped successively in abduction, and at a range of positions two scans were run. A 3D sequential acquisition at high spatial resolution (fast gradient echo sequence with radial reconstruction: FFE, TR/TE 6.4/3.1ms, Flip angle 15deg, FOV/matrix 500mm/410x512) was run to localize the hip position (gold standard) and secondly the optimized 2D dynamic protocol was run (seven imaging planes, gradient echo sequence with balanced gradients: bFFE, TR/TE 3.5/1.1ms, Flip angle 80deg, pixel size 2 x 2mm, slice thickness 10mm, partial Fourier reduction fac-

tor of 0.65 in read direction). The slice positions of the dynamic slices were required to be adjusted to intersect appropriate bony landmarks on each volunteer. These planes were set initially and maintained throughout the sequential motion protocol. In order to analyse optical markers/ bones relative displacement, markers, filled with contrast agent, were imaged at the same time in the scanner.

3.2. Bone motion tracking

In order to deduce kinematical properties of the musculoskeletal system, techniques have been developed to measure internal motion of organs. The use of bone screws or implantable markers provides a gold standard of bone motion measurement, although it is a very invasive approach. Nowadays, medical imaging technology has reached a level where it is possible to capture internal motion with different modalities (CT, MRI, US, SPECT). It opens up a way of measuring motion non-invasively with image-based methods. Depending on the dMRI protocol, between two and six planes can be acquired simultaneously, while preserving real-time aspect (>6 frames/sec). We present a new technique to track bones motion automatically from dynamic MRI based on the original combination of temporal information of dMRI and spatial information of static MRI.

Dynamic MRI is incomplete in terms of volumetric information, although it provides temporal information on single slices. On the other hand, static MRI provides prior knowledge on the bone morphology and is successfully segmented using the automatic deformable-model based approach presented before. In the static volume, normalized coordinate systems of the femur (Sf) and the pelvis (Sp) are oriented using anatomical landmarks and centred on the hip joint center (HJC), which is previously calculated with a functional method [KSM*02]. The problem of tracking bones in dynamic MRI is equivalent to a rigid registration problem (with six parameters) between the set of 2D dynamic images to the 3D static volume. Various registration methods have been proposed in the literature [Bro92]. They always correspond to a functional energy minimization problem. The

energy, calculated with a similarity metric Δ , measures degree of matching. For each bone, we automatically define a mask in the static volume from the bone model, where transformations are considered to be rigid (Figure 3). It contains the bone and surrounding tissues at a distance (determined empirically) that is less than 5mm. Δ is calculated only for the points located inside the mask. For the minimization process, we choose to use the six parameters describing the relative transformation between S_f and S_p in the dynamic acquisition coordinate system. After pelvis and femur tracking, we obtain normalized abduction, flexion and internal/external rotation angles and the displacement of the HJC obtained with the predictive method [KSMMT03] that uses surface models only. We have tested different similarity metrics based on a region or a frontier approach: normalized cross-correlation (NCC), absolute differences (AD), mutual information (MI), normalized mutual information (NMI) and model matching (MM). NNC, AD, MI and NMI are standard metrics [HHD*00]. MM aims at aligning the model with image edges in the dynamic slices. NCC, AD and MM can be applied to the gradient vector images. It provides a better robustness to rotation tracking. Information in the static volume, where positions of the transformed dynamic slices voxels are floating, is trilinearly interpolated. Optical Markers are tracked automatically on images using, similarly to bones, a template matching method. In this case, the reference volume is located at the first frame where landmarks have been manually initialized.

After the coarse manual initialization, the amoeba optimizer [NM65] is used to minimize the similarity measure automatically. We validated the technique by comparing, on 6 different postures and 6 healthy volunteers, relative positions between femur and pelvis, assuming the volunteer remained in exactly the same position for both high-resolution sequential MRI, providing a gold standard bones position measurement, and the dynamic MRI acquisition. The error is the modulus of the translation and rotation transformation vectors between the two positions. We compared similarity metrics in terms of tracking error and robustness around the final solution in order to optimize bone motion estimation. Subsequently, we compared results for various plane numbers, positions and resolutions in order to optimize the acquisition in terms of speed. Normalized cross-correlation based on gradient images gave the most accurate tracking and was the more robust metric. We optimized the number of acquisition planes to three along with the definition of their optimal position and orientation. We found that decreasing resolution down to 4x4mm could improve acquisition speed preserving an acceptable tracking error. The final error was $3.3\bar{f}$ ($\text{dev}=1.5\bar{f}$) in terms of relative position between the femur and the pelvis and the final frame rate for the dynamic protocol was 6.7 frames/sec.

4. Optimizing the optical markers configuration

In order to optimize the markers configuration, we use the bone motion tracking technique described above and we proceed as follows: Nine reflective markers are injected with a contrast agent and placed at specific anatomical locations on the right limb of the volunteer. The volunteer undergoes the MRI scanner; his/her right limb is constrained by a device and imaged by the scanner during each frame. The MR series are processed and the trajectories of the visible markers are calculated using the automatic tracking method detailed in the previous section. Therefore, each marker m_i is associated with an error r_i corresponding to the sum of its displacements from frame to frame as in the following:

```

For each marker  $m_i$ 
  For each frame  $f_i$ 
    Calculate the marker's sum of displacements  $\sum d_i$ 
  Assign to  $m_i$  the error  $r_i = \sum d_i$ 

For each triplet of marker ( $m_i, m_j, m_k$ )
  If ( $\text{Check\_colinearity}(m_i, m_j, m_k)$ )
    Calculate the distance  $d_{ijk} = d(m_i, m_j) + d(m_i, m_k) + d(m_j, m_k)$ 
    Calculate the error  $r_{ijk} = r_i + r_j + r_k$ 

Check_colinearity ( $m_i, m_j, m_k$ ): returns yes if the three markers  $m_i, m_j$  and  $m_k$ 
are non-collinear and no otherwise.

```

The best three markers m_i, m_j, m_k are the most distant ones (maximum value of the sum of the markers inter-distances d_{ijk}) with the less relative motion (minimum value of the sum of the markers displacements r_{ijk}). Thus, each triplet of markers is assigned two weights: r_{ijk} and d_{ijk} . To determine the best three markers, we need to minimize the quantity r_{ijk} and maximize the quantity d_{ijk} . [YCGMMT04]. In other words, we seek for the triplet that maximizes the fraction d_{ijk}/r_{ijk} . An exhaustive search is used and the best triplet is selected (Figure 4). The procedure is applied to three motion patterns recommended by our medical partner as being the clinical movements used to determine the range of motion of the patient: (hip abduction, flexion and rotation).

5. Motion capture using the optical system

5.1. Subject-fair visualization model creation

The reconstructed bone surfaces are simplified in lower polygon count to allow real-time display. They are inserted in a virtual human skin surface generated from adaptation of a generic model according to manual measurements of the subject's segments [SMT03]. Additionally the high-resolution reconstructed bones are used for evaluating the hip joint centre position in a dedicated application [KSMMT03]. This hip joint centre (HJC) is then set on the subject's technical skeleton model. This step ensures that the model's HJC matches the precisely evaluated HJC therefore providing realistic animation visualization. As a result we obtain a visualization model composed of fairly accurate hip joints with bones within an approximated subject's



Figure 3: A dynamic sample slice with its corresponding interpolated image in the static volume, masked images and 3D position in the static volume.

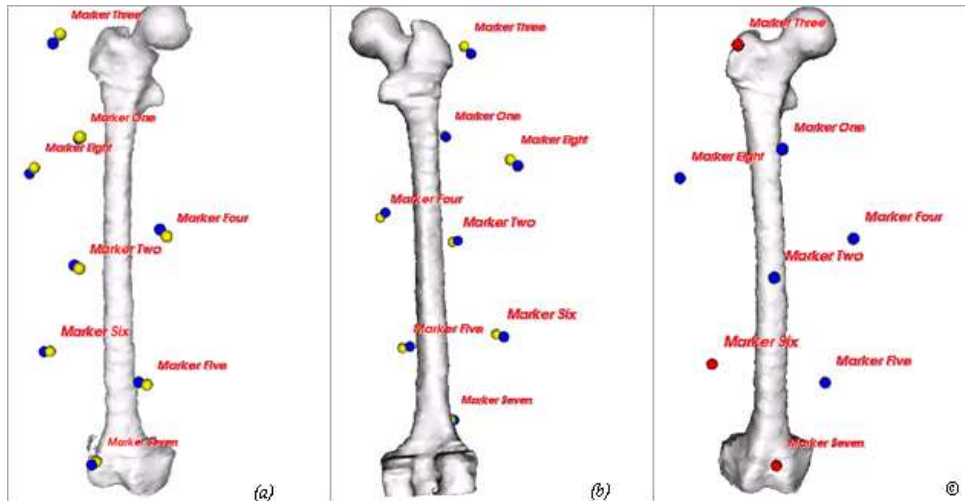


Figure 4: (a)(b) markers displacements: Yellow= real markers positions, Blue=calculated positions (c): Best markers selection (in red).

body deformable envelope (Figure 5) One key interest of our approach lies in the fact that the visualization model corresponds to the subject's real anatomy in the focus area (e.g. hip joint). The visualization of the motion that is as well recorded from the subject herself/himself is therefore closer to the real situation. That way, we removed the mapping on a subject-unrelated model bottleneck that gives little confidence in the visualization process given the variability in anatomy among different subjects.

5.2. Subject-fair external motion capture

For recording of subject movements, we use an optical motion capture system composed of 8 video cameras. The reflective skin markers are placed on anatomical landmarks of the subject according to the optimized configuration. The recorded markers trajectories are then converted into the joint space parameters of the subject's model. The converter technique [MBT99] takes into account the geometry of the skeleton model, motivating further the accurately matched subject to model process. A record of the subject in stand-up calibration posture is used as a subject/model posture map-

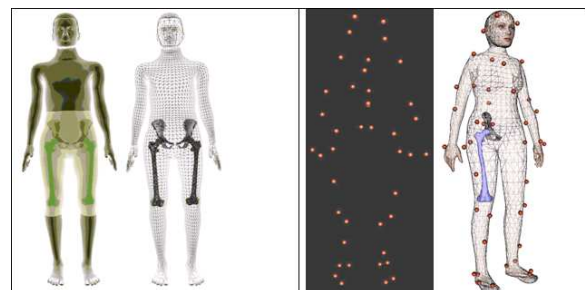


Figure 5: (a) Male subject model (volunteer) with hip reconstructed bones and body-sized skin (b) Markers and model in stand-up calibration posture after fine-tuning registration

ping reference. The model posture can be fine tuned with respects to the subject's recorded posture before converting the trajectories into animation. This is done in practice by visualization of the markers position in the stand-up posture and adjusting the model posture, thus creating an offset posture (Figure 6). This offset posture is then used in place of

the model's default posture in the process. Unlike MRI scanner, the optical system allows the recording of large range motions (Figure 6) that are typically used for detection of impingements. We record hip abduction/adduction (Figure 7), rotation, flexion, and conical motion. The positions of optical markers are visualized at the same time as the model during its animation; this serves as reference to assess the reliability of the animation mapping. Although we are concerned primarily by the study of the hip joints, we believe that providing a more complete, yet less accurate, animated visualization of the rest of the body is desirable as it confers a panel of views from general to detail. Similarly while the skinning of the model is purely geometrical, it is less abstract than pure optical markers and bones display.

The real-time visualization application we have developed is based on the VHD++ middleware framework [PPM*03]. For the purpose of this work, we integrated the management of optical markers animation and enhanced the virtual human production pipeline to satisfy the constraints of anatomical accuracy.

6. Conclusion

We have presented a methodology for patients' hip bones modelling and motion visualization using non invasive approaches. This technique improves the estimation of bones position in motion capture by using an optimized markers configuration to reduce artefacts due to skin and fat deformation. The markers configuration optimization is carried out using multi-slice dynamic MRI where bones and markers positions are tracked with an automatic and optimized tracking algorithm. Visual comparisons with classical animation based on common markers positions and standard animation skeleton showed an improvement in the animation realism. In the future, we plan to investigate methods to validate our technique. Existing techniques are invasive and therefore, of limited use. Single-plane video-fluoroscopy offers the best compromise between invasiveness and image quality and seems to be the best alternative. In addition, soft-tissue modeling is under investigation inside the framework of our project. They will be integrated into the motion visualization framework in order to express relevant information for clinical diagnosis (i.e. cartilage stress and strain). This methodology will be applied to the knee and ankle in order to obtain individualized and animated models of the full leg.

7. Acknowledgements

This work is supported by CO-ME (Computer Aided and Image Guided Medical Interventions) project funded by Swiss National Research Foundation. We would like to thank Dr. J.P. Vallée and Dr. H. Sadri from the Hopitaux Universitaires de Genève for their collaboration.

References

- [Bro92] BROWN L.: A survey of image registration techniques. *ACM Computing Surveys* 24, 4 (Dec. 1992), 325–376. 2, 3
- [HHD*00] HOLDEN M., HILL D., DENTON E., JAROSZ J., COX T., ROHLFING T., GOODEY J., HAWKES D.: Voxel similarity measures for 3-d serial mr brain image registration. *IEEE Transactions on Medical Imaging* 19, 2 (Feb. 2000), 94–102. 4
- [KSM*02] KANG M., SADRI H., MOCCOZET L., MAGNENAT-THALMANN N., HOFFMEYER P.: Accurate simulation of hip joint range of motion. *Proceedings of IEEE Computer Animation* (2002), 215–219. 1, 3
- [KSMMT03] KANG M., SADRI H., MOCCOZET L., MAGNENAT-THALMANN N.: Hip joint modeling for the control of the joint center and the range of motions. *Proceedings of the IFAC symposium on modelling and control in biomedical systems* (Aug. 2003), 23–27. 4
- [KWT88] KASS M., WITKIN A., TERZOPOULOS D.: Snakes: Active contour models. *International Journal of Computer Vision* 1 (1988), 321–331. 2
- [LC87] LORENSEN W., CLINE H.: Marching cubes: A high resolution 3d surface construction algorithm. *Computer graphics, Siggraph* 21, 39 (1987), 163–169. 2
- [LM01] LÖTJÖNEN J., MÄKELÄ T.: Segmentation of mr images using deformable models: Application to cardiac images. *International Journal of Bioelectromagnetism* 3, 2 (2001), 37–45. 2
- [MBT99] MOLET T., BOULIC R., THALMANN D.: Human motion capture driven by orientation measurements. *Presence, MIT* 8, 2 (1999), 187–203. 5
- [MTKG02] MAGNENAT-THALMANN N., KANG M., GOTO T.: Problems and solutions for the accurate 3d functional modelling of the hip and shoulder. *Proceedings of IEEE Computer Graphics International Organized by Computer Graphics Society (CGS)* (2002), 3–23. 1
- [NM65] NELDER J., MEAD R.: A simplex method for function minimization. *Computer Journal* 7 (1965), 308–313. 4

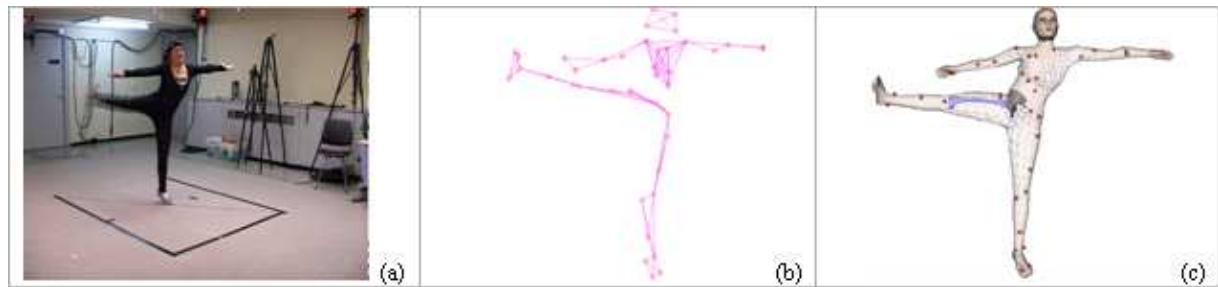


Figure 6: (a) Woman subject motion capture (b) Markers labeling (c) Model motion mapping



Figure 7: Large hip range of motion visualization example: abduction/adduction

- [PPM*03] PONDER M., PAPAGIANNAKIS G., MOLET T., MAGNENAT-THALMANN N., THALMANN D.: Vhd++ development framework: Towards extendible, component based vr/ar simulation engine featuring advanced virtual character technologies. *Computer Graphics International (CGI)* (2003). 6
- [RSS*96] ROHR K., STIEHL H. S., SPRENGEL R., BEIL W., BUZUG T. M., WEESE J., KUHN M. H.: Point-based elastic registration of medical image data using approximating thin-plate splines. *Visualization in Biomedical Computing* (1996), 297–306. 2
- [SMT03] SEO H., MAGNENAT-THALMANN N.: An automatic modeling of human bodies from sizing parameters. *ACM SIGGRAPH 2003 Symposium on Interactive 3D Graphics* (2003), 19–26. 4
- [SZL92] SCHROEDER W., ZAGE J., LORENSEN W.: Decimation of triangle meshes. *Siggraph 26, 2* (1992), 65–70. 2
- [vtk] www.vtk.org. 2
- [YCGMMT04] YAHIA-CHERIF L., GILLES B., MOLET T., MAGNENAT-THALMANN N.: Motion simulation of the hip joint using an optimized markers configuration. *Proceedings of the Eighth International Symposium on the 3-D Analysis of Human Movement* (2004), 57–60. 4