Human Activity Modeling and Simulation with High Biofidelity

Zhiqing Cheng, Stephen Mosher
Jeanne Smith, and Isiah Davenport
Infoscitex Corporation
Dayton, Ohio, USA
{zcheng, smosher}@infoscitex.com
{jsmith, idavenport}@infoscitex.com

John Camp and Darrell Lochtefeld
711th Human Performance Wing
Air Force Research Laboratory
Dayton, Ohio, USA
john.camp@wpafb.af.mil
darrell.lochtefelt@wpafb.af.mil

ABSTRACT

Human activity Modeling and Simulation (M&S) plays an important role in simulation-based training and Virtual Reality (VR). However, human activity M&S technology currently used in various simulation-based training tools and VR systems lacks sufficient biofidelity and thus is not able to describe and demonstrate the nuances of human activities and human signatures. This inadequacy becomes crucial when the training or the use of VR is human centered, such as human threat recognition training and dismount detection training. Human signatures that can be observed from a fairly long distance include body shape, gesture, and motion. In recent years, the Air Force Research Laboratory has investigated human modeling and simulation with high biofidelity, with an emphasis on true human shape and motion. This paper presents the technical development from these investigations, which include (a) static shape modeling and morphing; (b) pose modeling and dynamic modeling; (c) motion capture (in particular, markerless motion capture); (d) inverse kinematics and motion mapping/creation; and (e) creation and replication of human activity in 3-D space with true shape and motion. A brief review is conducted to discuss the methods and techniques related to these topics, along with some research results. Examples are provided to illustrate the importance of biofidelity in the simulation-based training.

ABOUT THE AUTHORS

Dr. **Zhiqing Cheng** is a principal engineer of Infoscitex Corporation, providing support to the Air Force Research Laboratory (AFRL) for the programs of human identification and activity recognition based on human bio-signatures. He has vast research experience in the areas of human modeling and simulation, artificial intelligence and machine learning, computer vision, and optimization. He has assumed many R&D projects as the principal investigator or as a major investigator and has published over 70 technical papers as the lead author. Currently, as a technical lead, he is working on the HMASINT program, a multi-directorate endeavor by the AFRL, to develop technologies for human measurement and intelligence information.

Dr. **John Camp** is a computer research scientist employed by the United States Air Force Research Laboratory (AFRL). Dr. Camp is the Government lead for AFRL's biofidelic human avatar program. As a program manager, he has been responsible for the successful development and implementation of a wide variety of products representing virtual humans in modeling and simulation. His research interests include computer vision and graphics. Dr. Camp received a B.S. in mathematics from the University of Florida, a Masters in computer systems from the Air Force Institute of Technology, and a Ph.D. in Computer Engineering from Wright State University.

Dr. **Darrell Lochtefeld** is a senior operations research analyst employed by the United States Air Force Research Laboratory (AFRL). As a scientist for the H-MASINT program, Dr. Lochtefeld directs and executes basic and applied research related to the discovery and exploitation of human biosignatures. His research interests include heuristic optimization, computational intelligence, data mining, modeling and simulation, and computer vision. Dr. Lochtefeld is the author of numerous conference proceedings, technical papers, and journal papers including research appearing in the Journal of Soft Computing and IEEE Transactions on Evolutionary Computation. Dr. Lochtefeld received a B.S. in computer science from Wright State University, a Masters in Business Administration (MBA) from the Ohio State University, and a Ph.D. in engineering from Wright State University.

Human Activity Modeling and Simulation with High Biofidelity

Zhiqing Cheng, Stephen Mosher
Jeanne Smith, and Isiah Davenport
Infoscitex Corporation
Dayton, Ohio, USA
{zcheng, smosher}@infoscitex.com
{jsmith, idavenport}@infoscitex.com

John Camp and Darrell Lochtefeld
711th Human Performance Wing
Air Force Research Laboratory
Dayton, Ohio, USA
john.camp@wpafb.af.mil
darrell.lochtefelt@wpafb.af.mil

INTRODUCTION

Human activity modeling and simulation (M&S) plays an important role in simulation-based training and virtual reality (VR). However, the human activity M&S technology currently used in most simulation-based training tools and VR systems lacks sufficient realism. In order to virtually describe and demonstrate the nuances of human activities and human signatures, modeling human shape and motion with high biofidelity is crucial.

Using conventional human modeling tools (e.g., Blender, 3dsMax, and Maya) or game engines (e.g., CryEngine 3, VBS2, and Delta3D), human activity modeling includes character building that creates its shape model and character animation that drives the model with the prescribed motion, both of which are associated with a skeleton model of the character. The shape model is defined by the surfaces attached to the skeleton, and the process of attaching surfaces to the skeleton is usually called skinning. The prescribed motion is given by the gross motion (translation and rotation of the whole body) and a sequence of poses that in turn, is defined by the joint angles for each pose. As the skeleton is driven by the prescribed motion, the attached surfaces will move accordingly and deform in a certain pattern which is controlled by specific blending schemes of the tools used. Therefore, in order to achieve high biofidelity for human activity M&S, it is essential to attain high biofidelity in the M&S of human shape and motion.

From the perspective of the motion status of a subject to be modeled, human shape modeling can be classified as either static or dynamic. Static shape modeling creates a model to describe the human shape at a particular pose, usually a standing pose. The static model can be used for human activity modeling as a character shape model. Dynamic shape modeling deals with shape variations due to pose changes or due to the subject being in motion. Apart from conventional approaches for human activity modeling and simulation, dynamic shape modeling has emerged as a viable alternative technique and shown its great potential for human activity modeling. Dynamic human shape modeling describes human shape changes during motion and thus can be used to directly replicate human activities in a 3-D space.

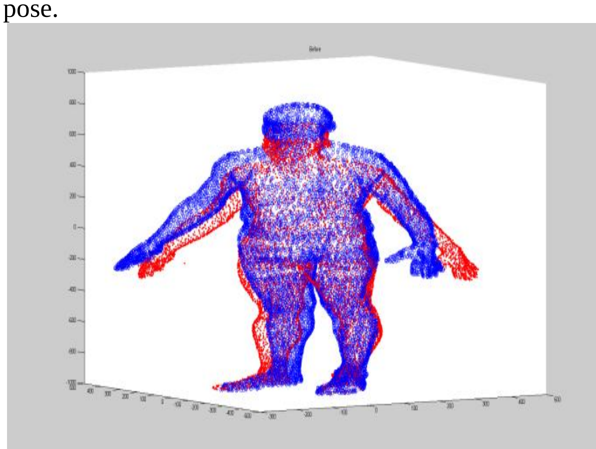
In recent years, a series of research activities has been performed at the Air Force Research Laboratory on human modeling and simulation, with an emphasis on high biofidelity and the goal to recognize human activities. This paper presents the results of these studies, along with discussions on the topics of static and dynamic human shape modeling, human motion capture and creation, and human activity replication and creation.

STATIC HUMAN SHAPE MODELING

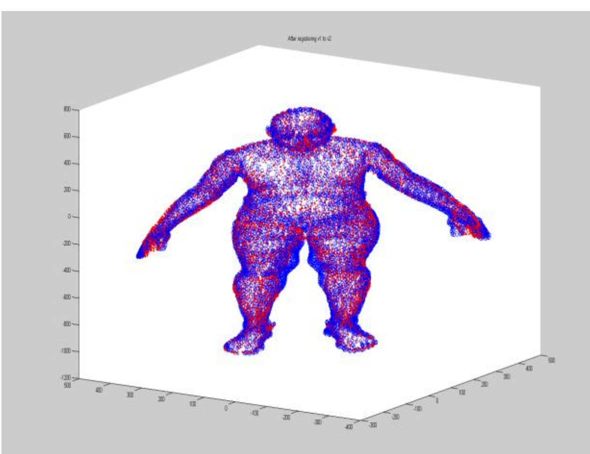
Software tools such as MakeHuman (http://www.makehuman.org/, a free software tool) are now available to create various generic human shape models with input parameters for gender, height, weight, etc. While these human shape models provide a realistic, graphical description of human body shape, they are often not able to depict the unique features that are associated with an individual or with a particular racial or ethnic group and thus lack the desired biofidelity. With advances in surface digitization technology, a 3-D surface scan of the whole body can be acquired in a few seconds. Whole body 3-D surface scans provide a very detailed capture of the body shape. Based on body scan data, human shape modeling with high biofidelity becomes possible. However, scan data files are usually very large and noisy and require further processing before becoming usable for shape modeling. The major issues involved with static shape modeling using scan data include surface registration, shape variation characterization, and shape reconstruction.

Surface Registration

Surface registration or point-to-point correspondence among the scan data of different subjects is essential to many problems of human shape modeling, such as shape parameterization and characterization, human shape variability (Allen et al., 2003; Azouz et al., 2005), and pose modeling and animation (Allen et al., 2002; Anguelov et al., 2005) where multiple subjects or multiple poses are involved. The method used for surface registration in this paper is called Coherent Point Drift (CPD), which can be used to register two point sets rigidly or non-rigidly. The description of CPD is provided in (Myronenko and Song, 2010). Often, the number of surfaces (accordingly the number of points) of the original scan data may be too large to be handled by the available computer memory on a typical workstation. Also, the original data may contain poorly formed polygons, webs between adjacent surfaces such as fingers and holes in the mesh. Therefore, the original scan data were smoothed and then simplified. After the number of faces was reduced to 20,000, the registration process was successfully completed. Figures 1 (a) and (b) illustrate the registration results of two different subjects in the same pose.



(a) Before registration



(b) After Registration

Figure 1. Surface (point-to-point) registration between two different subjects in the same pose

Shape Variation Characterization

The human body comes in many shapes and sizes. Characterizing human shape variation is traditionally the subject of anthropometry—the study of human body measurement. The sparse measurements of traditional anthropometric shape characterization curtail its ability to capture the detailed shape variations needed for realism. While characterizing human shape variation based on a 3-D range scan could capture the details of shape variation, the method relies on three conditions: noise elimination, hole-filling and surface completion, and point-to-point correspondence. Also, whole body scanners generate large data files that cannot be used directly for shape variation analysis. Therefore, it is necessary to convert 3-D scans to a compact representation that retains information of the body shape. Principal components analysis (PCA) has often been used as a solution to the problem. Allen et al. (2003) captured the variability of human shape by performing PCA over the displacements of the points from the template surface to an instance surface. Anguelov et al. (2005) also used PCA to characterize the shape deformation and then used the principal components for shape completion. Ben Azouz et al. (2005) applied PCA to the volumetric models where the vector is formed by the signed distance from a voxel to the surface of the model.

Shape Parameterization

For human shape modeling, it is desirable to have a set of parameters to describe human shape and its variation among different subjects. Human body shape can be parameterized in three different levels.

- Using surface elements. After surface registration of scan data among all subjects, the same set of vertices or other surface elements can be used to describe different body shapes (3D surfaces) (Allen et al., 2003; Anguelov et al., 2005). In other words, different body shapes are parameterized by the same set of vertices. While this method of characterization usually incurs a large number of parameters, a body shape can be directly generated from these parameters.
- Using principal component coefficients. After PCA, human body shape space is characterized by principal components. Each shape can be projected onto the eigenspace formed by principal components. Within this space, a human shape can be parameterized by its projection coefficients (Allen et al., 2003; Azouz et al., 2005). If the full eigenspace is used, perfect reconstruction can be achieved from the parameters to the body shape.
- Using anthropometric features. The relationship between eigenvectors and human anthropometric

features (e.g., height and weight) can be established through regression analysis (Allen et al., 2003; Azouz et al., 2005), and then a body shape can be parameterized by these features. This type of parameterization is not an exact mapping between a human body shape and its anthropometric features. Perfect reconstruction of a body shape usually cannot be achieved given a limited number of features.

Shape Reconstruction

Given a number of scan data sets of different subjects, a novel human shape can be created that will have resemblance to the samples but is not the exact copy of any existing one. This can be realized in three ways.

- Interpolation or morphing. One shape can be gradually morphed to another by interpolating between their vertices or other graphic entities. In order to create a faithful intermediate shape between two individuals, it is critical that all features are well-aligned; otherwise, features will cross-fade instead of move. Figure 2 illustrates shape morphing from one male subject to a female subject performed by the authors (Cheng et al, 2009).
- Reconstruction from eigenspace. After PCA analysis, the features of sample shapes are characterized by eigenvectors or eigen-persons which form an eigenspace. Any new shape model can be generated from this space by combining a number of eigen-persons with appropriate weighting factors (Azouz et al., 2005).
- Feature-based synthesis. Once the relationship between human anthropometric features and eigenvectors is established, a new shape model can be constructed from the eigenspace with desired features by editing multiple correlated attributes, such as height and weight (Allen et al., 2003) or fat percentage and hip-to-waist ratio (Seo et al., 2003).

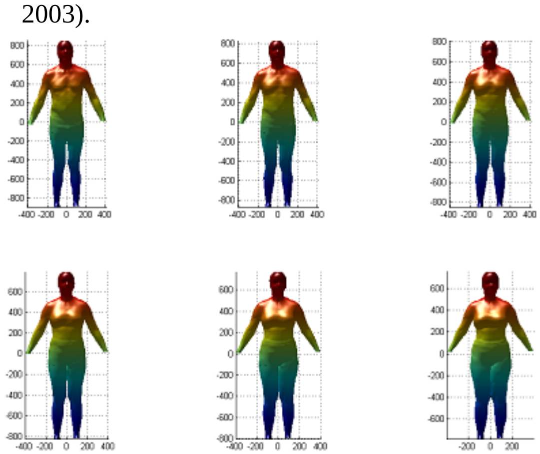


Figure 2. Morphing from one subject to another

DYNAMIC SHAPE MODELING

Dynamic shape modeling deals with shape variations due to pose changes or due to the subject being in motion. Two major issues involved in dynamic shape modeling are surface (shape) deformation with respect to pose changes and dynamic shape capture and reconstruction.

Body Deformation Modeling

Two main approaches for modeling body deformations are anatomical modeling and example-based modeling. The anatomical modeling is based on an accurate representation of the major bones, muscles, and other interior structures of the body (Aubel and Thalmann 2001). The finite element method is the primary modeling technique used for anatomical modeling. In the example-based approach, a model of some body part in several different poses with the same underlying mesh structure can be generated by an artist. These poses are correlated to various degrees of freedom, such as joint angles. Lewis et al. (2000) and Sloan et al. (2001) developed similar techniques for applying example-based approaches to meshes. Instead of using artist-generated models, recent work on the example-based modeling uses range-scan data. Allen et al. (2002 & 2003) presented an example-based method for calculating skeleton-driven body deformations. Their example data consists of range scans of a human body in a variety of poses. Using markers captured during range scanning, a kinematic skeleton is constructed first to identify the pose of each scan. Then a mutually consistent parameterization of all the scans is constructed using a posable subdivision surface template. Anguelov et al. (2005) developed a method that incorporates both articulated and non-rigid deformations. A pose deformation model was constructed from a training set of scan data that derives the non-rigid surface deformation as a function of the pose of the articulated skeleton. A separate model of shape variation was derived from the training data also. The two models were combined to produce a 3-D surface model with realistic muscle deformation for different people in different poses. The integrated model is called SCAPE (Shape Completion and Animation of People).

The method developed for pose deformation modeling in this paper employs the template model associated with the pose data set (Anguelov et al. 2005). It consists of 16 segments, each of which has the pre-defined surface division. The method consists of multiple steps, which are described below.

Coordinate Transformation

The body shape variations caused by pose changing and motion can be decomposed into rigid and non-rigid deformation. Rigid deformation is associated with the orientation and position of segments. Non-rigid deformation is related to the changes in shape of soft tissues associated with the segments in motion, which, however, excludes local deformation caused by muscle action alone. In the global (body) coordinate system, a segment surface has the articulated motion and surface deformation. However, in the local (segment) coordinate system, a segment surface has deformation only. Therefore, by transforming the global coordinate system to the local system, the effect of the articulated motion on each segment could be eliminated.

Surface Deformation Characterization

Suppose the surface deformations of each segment are collected in all poses. Then PCA can be used to find the principal components of the surface deformation for each segment. Figure 3 illustrates the eigen value percentage ratio in each component (total 70) of all segments (total 16). It is shown that for all segments, the variance (eigen value ratio) of principal components increases sequentially, and significant principal components are those from the order of 60 to 70. As PCA exploits the underlying characteristics of a data set, the surface deformation of a segment in all observed poses can be characterized by these principal components. The surface deformation in a particular pose can be decomposed or projected in the space that is formed by the PCs. Each decomposition/projection coefficient represents the contribution or effect from the corresponding PC.

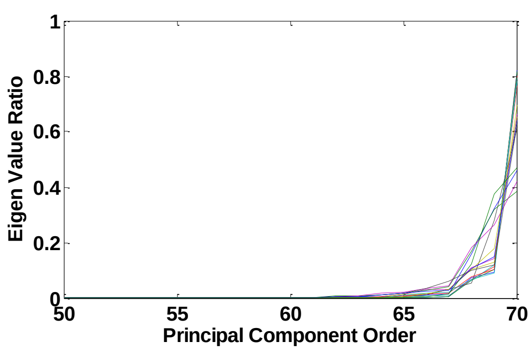
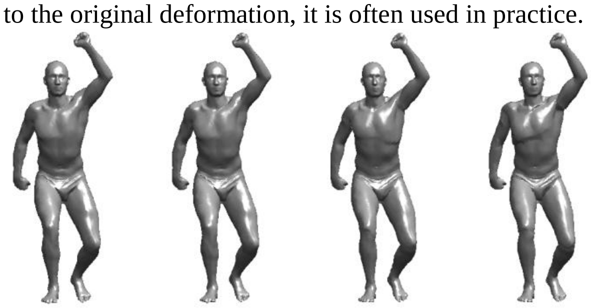


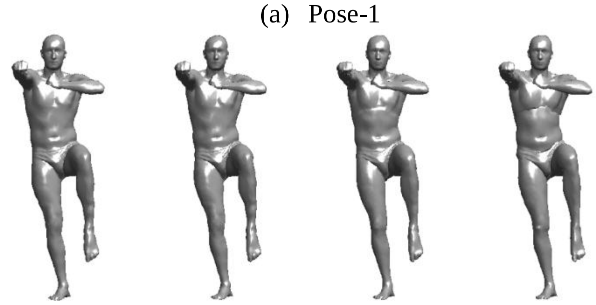
Figure 3. Eigen value ratio for all 16 segments.

Surface Deformation Reconstruction

The decomposition/projection coefficients can be used to reconstruct surface deformation. There are two types of reconstruction: (a) Full reconstruction, which uses all the PCs or eigenvectors; and (b) Partial reconstruction, which uses a number of significant PCs. Figure 4 illustrates the reconstructed shape for 2 different poses. In each row of Figure 4, the first is the original shape, the second is the shape from full reconstruction, and the third and fourth are the shapes from partial reconstruction with 20 and 10 largest PCs, respectively. It is shown that the full reconstruction can completely reconstruct the original surface deformation in all poses, which means it is a perfect reconstruction, and partial reconstruction can provide a reasonable approximation of the original shape. While full reconstruction provides complete reconstruction of the original deformation, it is not necessary in many cases. On the other hand, the accuracy of partial reconstruction can be controlled by selecting a proper number of significant PCs. As partial reconstruction provides a reasonable simplification or approximation to the original deformation, it is often used in practice.



(a) Pose-1



(b) Pose-2

Figure 4. Shape reconstruction using principal components (First column: original shape; second column: full reconstruction; Third column: partial reconstruction with 20 largest PCs; Fourth column partial reconstruction with 10 largest PCs).

Surface Deformation Representation

As the surface deformation of a segment is assumed to depend only on the rotation of the joint(s) connected, the relationship between the surface deformation and joint rotations has to be known. Joint rotations can be conveniently represented by their twist coordinates. The surface deformation can be compactly represented by its decomposition/projection coefficients. Ideally, the surface deformation can be expressed as a function of joint rotations. The relation between surface deformation and joint rotations can be linear or nonlinear. An appropriate function needs to be identified. The same function can be applied to all poses. Then, the measurement of surface deformation and joint rotations in all poses can be used to estimate the parameters of the function.

Surface Deformation Prediction

It is not feasible to measure the surface deformation of each segment for all possible poses, because the human body has a large number of degrees of freedom and can take virtually an infinite number of different poses. As a matter of fact, only a limited number of poses can be investigated in tests, but it is often required to predict surface deformation for new poses that have not been observed. Three methods can be used to predict surface deformation.

- Method-1: using principal components. Given the joint twist angles for a segment to define a particular pose, projection coefficients can be estimated. Using the full or a partial set of principal components, the surface deformation is reconstructed.
- Method-2: taking the nearest neighbor pose. Given the joint twist angles, find the nearest neighbor to the prescribed pose and take its surface deformation as an approximation. The neighborhood is measured in terms of the Euclidean distance between the joint twist angles for the two poses.
- Method-3: interpolating between two nearest neighbors. Given the joint twist angles, find two nearest neighbors to the prescribed pose. The pose deformation is determined by interpolating between the deformations of these two neighbor poses.

Figure 5 illustrates the predicted shape for 8 different poses using method-2.

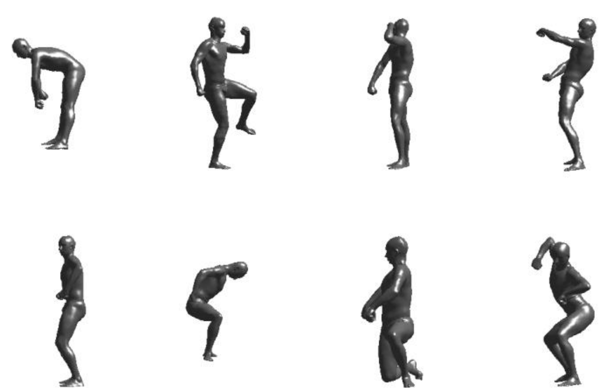


Figure 5. Predicted shape in 8 different poses.

Dynamic Shape Capture and Reconstruction

Dynamic Shape Capture

During dynamic activities, the surface of the human body moves in many subtle but visually significant ways: bending, bulging, jiggling, and stretching. Park and Hodgins (2006) developed a technique for capturing and animating those motions using a commercial motion capture system with approximately 350 markers. Supplemented with a detailed, actor specific surface model, the motion of the skin was then computed by segmenting the markers into the motion of a set of rigid parts and a residual deformation. Sand et al. (2003) developed a method (a needle model) for the acquisition of deformable human geometry from silhouettes. New technologies are emerging that can capture body shape and motion simultaneously at a fairly high frame rate (Nguyen and Wang, 2010; Izadi et al., 2011).

Shape Reconstruction from Imagery Data

- *From Photos*

Seo et al. (2006) presented a data-driven shape model for reconstructing human body models from one or more 2D photos. A data-driven, parameterized deformable model acquired from a collection of range scans of a real human body is used to complement the image-based reconstruction by leveraging the quality, shape, and statistical information accumulated from multiple shapes of range-scanned people. Guan et al. (2009) developed a method for estimating human body shape from a single photograph or painting.

- *From Video Sequences*

The recent work done by Balan et al. (2007) proposed a method for recovering human shape models directly from images. Specifically, the human body shape is represented by the SCAPE (Anguelov et al., 2005) and the parameters of the model are directly estimated from image data. A cost function between image observations and a hypothesized mesh is defined and the problem is formulated as an optimization. Hasler et al. (2009a) developed a method to estimate the detailed 3-D body shape of a person even if heavy or loose clothing is worn. Within a space of human shapes learned from a large database of registered body scans, the method fits a template model (a 3-D scan model of a person wearing clothes) to the silhouettes of video images using ICP (iterative closest point) registration and Laplacian mesh deformation.

HUMAN MOTION CAPTURE AND PREDICTION

Motion capture (mocap) technologies can be marker-based or vision-based. The challenges for motion analysis involve inverse kinematics (IK) and motion mapping and creation.

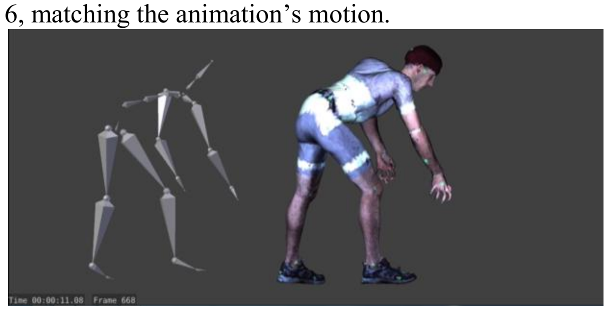
Marker-Based Motion Capture

As a traditional technique, marker-based motion capture technology has been developed to an advanced level that provides accurate and consistent measurements of body motion. The markers used in motion capture can be aligned with those used during body scanning thus providing some correspondence between body shape and skeleton motion. Various software tools are available for the analysis of motion capture data. The major limitations of marker-based motion capture technology include (a) it can only be used in a laboratory environment; (b) it has a limited coverage space; and (c) it requires subject cooperation. Several new technologies are emerging that use sensors mounted on the body (e.g., RF, accelerometers (Tautges et al., 2010), or mini-cameras (Shiratori et al., 2011)), enabling open-field motion capture.

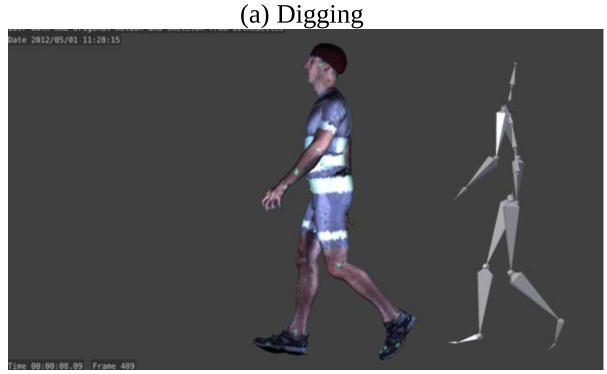
Markerless Motion Capture

As an active research area in computer vision for decades, markerless or vision-based human motion analysis has the potential to provide an inexpensive, unobtrusive solution for the estimation of body poses and motions. Extensive research efforts have been performed in this domain (Moeslund et al., 2006), which have been motivated by the fact that many application areas, including surveillance, human–computer interaction and automatic annotation, will benefit from a robust solution to the problem (Poppe 2007). Agarwal and Triggs (2006) developed a learning-based method for recovering 3-D human body pose from single images and monocular image sequences. Their approach requires neither an explicit body model nor prior labeling of body parts in the image. Instead, it recovers pose by direct nonlinear regression against shape descriptor vectors extracted automatically from image silhouettes. A recent development is capturing motion and dynamic body shape simultaneously from video imagery. Using SCAPE (Anguelov et al., 2005), Balan et al. (2007) developed a method for estimating the model parameters directly from image data. Their results showed that such a rich generative model as SCAPE enables the automatic recovery of detailed human shape and pose from images. Hasler et al. (2009b) presented an approach for markerless motion capture of articulated objects, which are recorded with multiple unsynchronized moving cameras. Instead of using fixed (and expensive) hardware synchronized cameras, their approach is able to track people with off-the-shelf handheld video cameras.

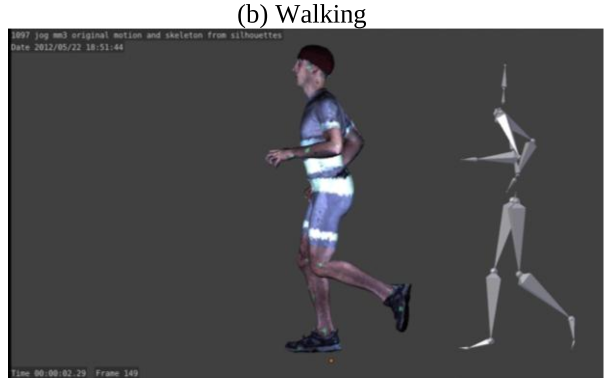
The approach developed by Agarwal and Triggs (2006) was implemented in this paper for markerless motion capture. As shown in Figure 6, using body scan and mocap data collected in the AFRL 3dHSL Lab, 3-D models were created for four activities (digging, walking, jogging, and throwing) using Blender (http://www.blender.org/). By animating the model of each activity, a sequence of 3-D shape models was generated for each activity, from which a sequence of silhouettes was derived. By establishing the relationship between image features (which are described by the histogram of shape context of silhouettes) and joint angles (which are used to define poses), the motion of the subject (which is defined by a sequence of poses) is captured. The resulting motion is applied to the skeleton shown in each image in Figure 6, matching the animation's motion.



(a) Digging



(b) Walking



(c) Jogging

(d) Throwing

Figure 6. Markerless motion capture from 2-D video imagery

Inverse kinematics

Inverse kinematics, the process of computing the pose of a human body from a set of constraints, is widely used in computer animation. However, the problem is often underdetermined. While many poses are possible, some poses are more likely than others. In general, the likelihood of poses depends on the body shape and style of the individual person. Grochow et al. (2004) developed an inverse kinematics system based on a learned model of human poses that can produce the most likely pose satisfying the prescribed constraints in real time. Training the model on different input data leads to different styles of IK. The model is represented as a probability distribution over the space of all possible poses. This means that the model can generate any pose, but prefers poses that are most similar to the space of poses in the training data. A common task of IK is to derive joint angles from markers, for which, OpenSim (https://simtk.org/home/opensim), an open source software package can be used.

Motion Mapping

Motion mapping and motion generation are two issues related to IK but have independent significance. It is desirable to map the motion from one subject to another, because it is not feasible to do motion capture for every subject and for every motion or activity. By assuming that different subjects will take the same key poses in an action or motion, one approach is mapping joint angles from one to another, as shown in Figure 7 where motion is mapped onto 3dsMax biped models. Note that since the pelvis is usually treated as the reference segment, the hip joint center vertical location needs to be adjusted to reflect the variation of subject size in order to ensure appropriate contact between the feet and the ground. While motion mapping may be fairly natural and realistic, it may not be able to provide sufficiently high biofidelity, because the differences between human bodies and the interaction between human body and boundaries are ignored.

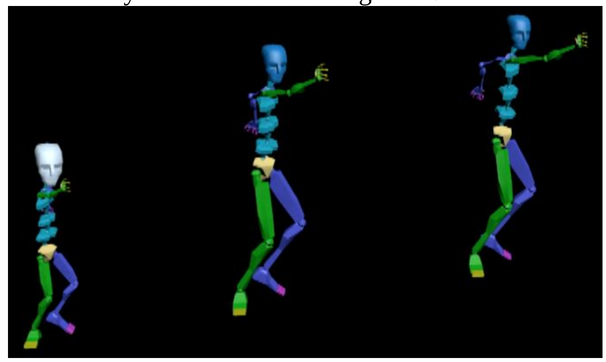


Figure 7. Mapping the captured motion into a group

Motion Creation

One method of motion creation is to create several key poses (frames) and then fills the gaps between those key poses via interpolation. This approach is often used by game developers. The created motion is based on human imagination and thus lacks realism and biofidelity, as shown in Figure 8. Alternatively, motion creation can be handled in more rigorous and scientific ways. Wei et al. (2011) showed how statistical motion priors can be combined seamlessly with physical constraints for human motion modeling and generation. The key idea of the approach is to learn a nonlinear probabilistic force field function from prerecorded motion data with Gaussian processes and combine it with physical constraints in a probabilistic framework. In addition, they showed how to effectively utilize the new model to generate a wide range of natural-looking motions that achieve the goals specified by users. Some tools were developed for motion creation based on biomechanics and physics, such as DANCE (http://www.arishapiro.com/), which is used for physics-based animation research, including dynamic simulation of rigid bodies, motion capture and dynamic control.

Figure 8. The comparison between two animations (mocap data vs. key framing data)

ACTIVITY REPLICATION AND CREATION

Replication

Activity replication is replicating a human activity that was recorded from a human subject in a laboratory using 3-D modeling. Technologies that are capable of capturing human motion and 3-D dynamic shapes of a subject during motion are not yet ready for practical use. Data that can be readily used for 3-D activity replication are not currently available. Alternatively, a motion capture system can be used to capture markers on the body during motion and a 3-D body scanner can be used to capture the body shape in a pose. Based on the body scan data and motion capture data, animation techniques can be used to build a digital model to replicate a human activity in 3-D space.

In this paper, open-source software was used for activity replication. MeshLab (http://meshlab.sourceforge.net/) was used to process 3-D scan data, OpenSim was used to derive skeleton models and the associated joint angles from motion capture data, and Blender was used to create an animation model that integrated body shape and motion. Human subject testing for data collection on human activities was conducted in the 3-D Human Signatures Laboratory (3DHSL) at the Air Force Research Laboratory (AFRL). The data collected included scans and mocap data.

The body scan data acquired consists of a large number of data points (vertices) (typically a half-million or more) and may contain holes and large openings. The data were processed so that it could be used for the modeling. MeshLab was used to clean-up the data and to fill holes. Smoothing and approximation functions in MeshLab were implemented to reduce the total number of faces for each subject scan to 50,000 and to create meshes of the body shape required for the modeling. OpenSim was used to derive a skeleton model from mocap data (TRC file) and to calculate the joint angles for the skeleton. The skeleton model and associated joint angles were put in a Bio-vision Hierarchical (BVH) file. Both the body surface mesh data and the BVH file were imported into Blender. Blender was used to integrate the shape with the motion and to create an animation model that replicates an activity. Figure 9 shows the models created for four activities (jogging, limping, shooting, and walking) at a particular frame. Note that activity replication can be done using commercial modeling tools (e.g., Autodesk 3dsMax and Maya).

Creation

Activity creation involves motion creation and dynamic shape creation. While some methods have been developed for motion creation, many issues remain. Creating a dynamic shape for any pose or activity is still a challenging task. Alternatively, in the following example, by matching body shape data with mocap data, two activities (diving-rolling and running-ducking) were created using body scan data and mocap data collected from different subjects. The mocap data for the two activities were derived from the Carnegie Mellon University (CMU) mocap database (http://mocap.cs.cmu.edu/). Using the lengths of major segments as the search criteria, the body shape data were derived from the CAESAR (Civilian American and European Surface Anthropometry Resource) database (Robinette et al., 1999). Then, 3-D animation models were created using Blender which fuses the shape and motion information together and deforms the body shape in accordance with body motion , as shown in Figure 10.

Figure 9. Replication of a subject in four activities: limping, jogging, shooting, and walking.

(a) Diving-rolling (b) Running-ducking

Figure 10. Activity creation using body scan data and mocap data from different subjects.

CONCLUSIONS

Biofidelity is a critical factor when human activity M&S is used in a virtual reality or training system that is human centered. In order to attain high biofidelity, a concerted effort for accurate human shape and motion data collection, motion analysis, and shape modeling must be undertaken. Based on subject tests and data collection, human activities can be replicated in 3-D space with fairly high biofidelity. The data-driven human activity models can be incorporated into highly fidelic 3-D scenario models to provide natural and realistic exposure and experience to trainees/users. However, it is not feasible to collect data for every subject and for every activity. Therefore, it is necessary to develop technologies for creating activities. Activity creation relies on dynamic shape modeling and motion creation, for which further investigations are needed to overcome remaining technical obstacles.

REFERENCES

Agarwal, A., & Triggs, B. (2006). Recovering 3-D Human Pose from Monocular Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No.1.

Allen, B., Curless, B., & Popovic, Z. (2002). Articulated Body Deformation from Range Scan Data. *ACM SIGGRAPH 2002*, 21-26.

Allen, B., Curless, B., & Popvic, Z. (2003). The space of human body shapes: reconstruction and parameterization from range scans. *ACM SIGGRAPH 2003*, 27-31.

Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., & Davis, J. (2005). SCAPE: Shape Completion and Animation of People, *ACM Transactions on Graphics (SIGGRAP 2005)*, 24(3).

Azouz, Z., Rioux, M., Shu, C., & Lepage, R. (2005). Characterizing Human Shape Variation Using 3-D Anthropometric Data. *International Journal of Computer Graphics*, volume 22, number 5, 302-314.

Aubel, A., & Thalmann, D. (2001). Interactive modeling of the human musculature. *Proc. of Computer Animation.*

Balan, A., Sigal, L., Black, M., Davis, J., & Haussecker, H. (2007). Detailed Human Shape and Pose from Images. *IEEE Conf. on Comp. Vision and Pattern Recognition (CVPR)*.

Cheng, Z., Portny, J., Smith, J., & Mosher, S. (2009). Dynamic 3-D Human Shape Modeling for Intention Prediction from Video Imagery. *SBIR Phase I report, AFRL-RH-WP-TR-2009-tr##*.

Grochow, K., Martin, S., Hertzmann, A., & Popovic, Z. (2004). Style-Based Inverse Kinematics. *Proc. of SIGGRAPH'04*.

Guan, P., Weiss, A., Balan, A., & Black, M. (2009). Estimating Human Shape and Pose from a Single Image. *Proceedings of ICCV*, 1381-1388.

Hasler, N., Rosenhahn, B., Thormahlen, T., Wand, M., Gall, J., & Seidel, H. (2009b). Markerless Motion Capture with Unsynchronized Moving Cameras. *IEEE Conference on Computer Vision Pattern Recognition, CVPR 2009*.

Hasler, N., Stoll, C., Rosenhahn, B., Thormahlen, T., & Seidel, H. (2009a). Estimating Body Shape of Dressed Humans. *Computers & Graphics* Vol 33, No 3, 211-216.

Izadi, S., Newcombe, R., Kim, D., Hilliges, O., Molyneaux, D., Hodges, S., Kohli, P., Shotton, J., Davison, A., & Fitzgibbon, A. (2011). KinectFusion: Real-Time Dynamic 3-D Surface Reconstruction and Interaction, a Technical talk on *Siggraph 2011*, Vancouver, Canada.

Lewis, J., Cordner, M., & Fong, N. (2000). Pose space deformations: A unified approach to shape interpolation and skeleton-driven deformation. *Proceedings of ACM SIGGRAPH 2000*, 165–172.

Moeslund, T., Hilton, A., & Krüger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, Vol. 104, No. 2-3, 90-126.

Myronenko, A., & Song, X. (2010). Point Set Registration: Coherent Point Drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 12.

Nguyen, D., & Wang, Z. (2010). High Speed 3-D Shape and Motion Capturing System. a Poster of *Siggraph 2010*, Los Angles, USA.

Park, S., & Hodgins, J. (2006). Capturing and Animating Skin Deformation in Human Motion. *ACM Transaction on Graphics (SIGGRAPH 2006)*, 25(3), 881-889.

Poppe, R. (2007). Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding*, Vol. 108, pp. 4-18.

Sand, P., McMilla, L., & Popovic, J., Continuous Capture of Skin Deformation. *ACM Transactions on Graphics* 22 (3), 578-586.

Seo, H., Cordier, F., & Thalmann, N. (2003). Synthesizing Animatable Body Models with Parameterized Shape Modifications. *Eurographics/SIGGRAPH Symposium on Computer Animation 2003*.

Seo, H., Yeo, Y., & Wohn, K. (2006). 3-D Body Reconstruction from Photos Based on Range Scan. Lecture Notes in Computer Science, 2006, No. 3942, 849-860, Springer-Verlag.

Shiratori, T., Parky, H., Sigal, L., Sheikhy, Y. & Hodginsy, J. (2011). Motion Capture from Body-Mounted Cameras. *ACM Trans. Graph. 30* (4), Article 31.

Sloan, P., Rose, C., & Cohen, M. (2001). Shape by example. *Proceedings of 2001 Symposium on Interactive 3-D Graphics*.

Tautges, J., Zinke, A., Krüger, B., Baumann, J., Weber, A., Helten, T., Müller, M., HSeidel, H., & Eberhardt, B. (2010). Motion Reconstruction Using Sparse Accelerometer Data. *ACM TOG* 30(3), article 18.

Wei, X., Min, J., & Chai, J. (2011). Physically Valid Statistical Models for Human Motion Generation. *ACM Trans. Graph.* Vol 30, No 3, Article 19.

Robinette, K., Daanen, H., and Paquet, E. (1999). The Caesar Project: A 3-D Surface Anthropometry Survey. *Second International Conference on 3-D Digital Imaging and Modeling (3DIM'99)*, 380–386, Ottawa, Canada.