# Dynamic Superimposition of Synthetic Objects on Rigid and Simple-deformable Real Objects

Yann Argotti†, Larry Davis†, Valerie Outters‡, and Jannick P. Rolland†‡

School of Electrical Engineering and Computer Science†
School of Optics/CREOL‡
University of Central Florida
4000 Central Florida Blvd
Orlando, FL 32816-2700

## Abstract

*A current challenge in augmented reality applications is the accurate superimposition of synthetic objects on real objects within the environment. This challenge is heightened when the real objects are in motion and/or are non-rigid. In this article, we present a robust method for real-time, optical superimposition of synthetic objects on dynamic rigid and simple-deformable real objects. Moreover, we illustrate this general method with the VRDA Tool, a medical education application related to the visualization of internal human knee joint anatomy on a real human knee.*

## 1. Introduction

A significant challenge in AR applications is the correct superimposition of synthetic objects on real objects within the environment. Real and synthetic objects must be placed into register, that is, spatial coincidence, from a common reference. The superimposition becomes more challenging when the real objects are moving. In general, real objects are considered rigid with respect to tracking, but this represents only a subset of the possibilities; the real objects in the environment may be non-rigid.

The contribution of this article is to present a method for dynamic superimposition that is robust, accurate, interactive-speed, and applicable to rigid and simple-deformable real objects. The method presented is applied to the Virtual Reality Dynamic Anatomy (VRDA) Tool, a visualization system developed for the study of complex joint motions [27]. Furthermore, the method presented here is applicable to other areas of AR, including surgical assistance, engineering applications, military simulation, and entertainment.

In the following paragraphs, we discuss related work in augmented reality, tracking, and anatomical motion tracking. We summarize the calibration technique associated with the dynamic superimposition [3] and focus the contribution of the paper on the real-time, dynamic superimposition method. Finally, we demonstrate the method as it applies to the VRDA Tool and present visual results of the superimposition. The dynamic superimposition is being assessed in the context of orthopedic research.

## 2. Previous Work

In superimposing synthetic objects on real objects, superimposition may occur statically (objects for superimposition are relatively still) or dynamically. Among the research in static superimposition methods for AR, Mellor introduced a static, marker-based method for tracking that was able to recover depth information using a single video source [15]. Grimson et al created an AR system that featured interactive updates of a patient's brain using MRI data [13]. Fuchs et al developed a system to aid in laparoscopic surgery that used structured light patterns for tracking [12].

Furthermore, dynamic superimposition within AR applications has been typically limited to rigid objects. Bajura and Neumann used a closed loop registration correction method to enhance dynamic superimposition [7]. Uenohara and Kanade implemented a method to dynamically track rigid objects using the video outputs of cameras in a video see-through HMD [24]. State et al implemented an occlusion-resistant, hybrid tracking scheme to achieve dynamic superimposition [22]. In the realm of wearable computing, Starner et al created a system to aid in everyday tasks [21], while Billinghurst and Kato developed a system that used a dynamic, collaborative writing surface [8].

In recent years, however, there has been increased interest in tracking the motion of non-rigid objects. Halvey and Weinshall implemented a method for tracking non-rigid objects in video sequences based upon optical flow methods [14]. Comaniciu et al also implemented a method for tracking non-rigid objects based upon statistical properties [11]. Within the context of tracking non-rigid, dynamic objects, one of the most challenging tracking tasks is tracking the motion of human anatomy. Spoor and Veldpaus published a method for calculating rigid body motion from the spatial coordinates of markers that has been adapted to tracking skeletal motion [20]. In addition, techniques have been devised that address the problems associated with accurately tracking anatomical motion [2][10].

In this paper, we examine the problem of tracking simple-deformable bodies within an augmented reality system and present a general method for dynamically superimposing synthetic objects on these real objects at interactive speeds.

## 3. Method Overview

The dynamic superimposition procedure assumes the use of a marker-based tracking system capable of providing the 3D location of the markers. Each real object in the system is then defined by a cluster of markers placed on its surface. The procedure also assumes the use of a stereoscopic display device with markers attached to it for determining the viewpoint of the user.

In the method presented, we call the tracker coordinate system the global coordinate system or global frame. Moreover, for each object in the environment (real or synthetic), we associate a local coordinate system or local frame. We refer to the transformation matrices between coordinate systems within the environment as links. Objects that have an expressed transformational relationship between one another are referred to as linked objects.

Simple-deformable objects are defined as objects that are slightly changing in shape compared to an equivalent rigid object. The change in shape can be quantified by the change in the eigenvalues of the dispersion matrix associated with a cluster of markers placed on the object [3] [4].

The first step in the dynamic superimposition procedure is to measure the global location of at least three markers on each real object. Given the location of these markers in their local coordinate frame, an optimization method based on singular value decomposition (SVD) is applied to estimate the rotation and translation, which when applied to the local coordinates, yield the measured global coordinates. The local motion of markers on a semi-deformable object is managed during this step. Next, because there may be a need for collision detection and/or motion constraints between linked objects within the environment, the transformation matrix which links the real objects is used as an input to a kinematic model of motion [6]. The last step is the stereoscopic rendering process that combines all the required transformation matrices and defines the relationship between the real and synthetic environment. Included in the final step is the correction of optical distortion that may be introduced by the display device.

## 4 Dynamic Superimposition Process

In this section, we detail each step of the dynamic superimposition algorithm. The superimposition method is robust, taking into account noise in the tracking data. The superimposition requires knowledge of the display device parameters (field of view, interpupillary distance, etc.). Also, the method requires the local coordinates for each marker and the links between all real objects and their synthetic counterparts, which are determined from a calibration procedure [3].

### 4.1. Locating Real Objects with All Markers Visible

To properly display synthetic objects from the eye viewpoint, the matrix describing the transformation from the real object local frame to the global frame is now needed.

Given $n$ markers on the surface of the real object, if $x_i$ is the $i^{th}$ real object marker coordinate expressed in the real object local frame and $y_i$ is the $i^{th}$ real object marker coordinate in the global frame, then the desired transformation best fits $x_i$ into $y_i$. The transformation can be decomposed into a rotation matrix, $R$, and a translation vector, $T$, such that $y_i = Rx_i + T, i \in [1, n]$. Scaling is unnecessary because the transformation is between normalized frames.

Furthermore, the data from the tracking system are intrinsically noisy and, in the case of a simple-deformable object, the markers may move with respect to each other. To take into account both the noise and the possible relative motion of the markers, a weight is applied to each marker to represent its fidelity. In this way, more importance is given to the higher fidelity markers.

The relative motion of a marker within the global frame is computed with an iterative process. The initial positions of the markers in the local frame are determined in the calibration procedure. The first time $R$ and $T$ are estimated, the initial marker locations are used as inputs to the optimization process. The local marker positions are then updated by using the $R$ and $T$ most recently obtained. The updated local marker locations, are then used to determine the relative displacement, $\Delta x$, expressed as the difference between the updated local marker locations and the original local marker locations. The standard deviation of $\Delta x$ measured over at least 10 frames of standard real object motion gives the relative motion of the markers.

The error, or sensitivity, of a marker location is determined by comparing the tracking system precision to the current marker relative motion and taking the larger value. Thus, the weight of the $i^{th}$ marker, $w_i$ is then given by

$$w_i = 1 - \frac{Error_i}{\sum_{i=0}^{n} Error_i} \quad , \tag{1}$$

where $n$ is the total number of markers.

As a result of the noise in the tracking data and the motion of the markers, approximations of $R$ and $T$ are estimated using least-squares minimization. The error, $e$, to minimize is given by

$$e(R,T) = \sum_{i=1}^{n} w_i \|y_i - Rx_i - T\|^2 \quad , \tag{2}$$

which can solved using SVD [23]. Once estimated, $R$ and $T$ are arranged in a 4x4 matrix format. We chose to implement a method based on SVD because this classical optimization method is robust [21], gives the best possible solution at all times, is computationally efficient [5], and converges quickly to a solution.

### 4.2 Locating Real Objects with Occluded Markers

During the tracking process, some markers may not be detected by the tracking system, while other markers are detected. Since we are weighting the marker coordinates to give more importance to the markers that are less perturbed by measurement noises and by relative motion on the real object, we now limit the computation to take into account only markers that have been detected. In this case, we have to compute the new weight, $w_i'$ associated to the $i^{th}$ marker whose current weight is $w_i$.

$${w_i}' = 1 - \frac{w_i}{\sum_{i=0}^{n} w_j \cdot Visible(j)} \quad , \tag{3}$$

$n$ is the total number of markers of the current real object and $Visible(j)$ is a function that gives 1 if the $j^{th}$ marker is detected, and 0 otherwise. Finally, we can determine the optimal solution for the rotation and translation using SVD, resulting in the transformation from the real object local frame to the global frame.

### 4.3 Computation of Relative Position and Orientation of Real Objects

In the previous sections, we described how to determine the transformation matrices, $M_{o\_oi}$, that relate the local coordinate system of the $i^{th}$ real object to the global coordinate system. We then determine the links between real objects in the environment by concatenating the transformations appropriately.

After computing the links between the objects, the transformation matrices describing the global position and orientation for each real object are used as entries in lookup tables. The lookup tables describe the motion between real objects and provide the location and orientation of their synthetic counterparts. The lookup table data are then used to modify the synthetic object transformation matrices to provide realistic, computationally efficient visualization.

### 4.4 Rendering Synthetic Objects

The last part of the dynamic superposition process is the rendering of synthetic objects for each eye view and the correction optical deformations introduced by the HMD optics [16][17].

Head tracking (or display device tracking) is utilized to determine the viewpoint of the user. We use the method in Section 4.1 to find $M_{h\_o}$, the transformation from the global frame, $o$, to the local head frame, $h$. To compute $M_{re\_sn}$ and $M_{le\_sn}$, the transformations from the $n^{th}$ synthetic object to the user's right eye and left eye viewpoints respectively, we can concatenate $M_{re\_h}$, the transformation from the head frame to the right eye frame and $M_{le\_h}$, the transformation from the head frame to the left eye frame with $M_{h\_sn}$, the transformation from the $n^{th}$ synthetic object to the local head frame.

$M_{h\_sn}$ is computed by concatenating $M_{on\_sn}$, the transformation from the $n^{th}$ synthetic object frame to the $n^{th}$ real object frame, $M_{o\_on}$, the transformation matrix from the $n^{th}$ real object frame to the global frame, and $M_{h\_o}$. We note that for both viewpoint transformations, there is a common part that represents the transformation from the $n^{th}$ synthetic object to the local head frame. Thus, we compute $M_{h\_sn}$ only once and then apply it to $M_{re\_h}$ and $M_{le\_h}$. As a result, any point defined in a synthetic object frame is transformed and expressed in each eye local frame. In OpenGL, we set the modelview matrix to $M_{sn\_re}$ and $M_{sn\_le}$ before rendering the $n^{th}$ synthetic object for the right and left viewpoints, respectively. To correct the residual optical distortion present in the display device, we apply the rendered image on a deformed polygon mesh by texture mapping [1].

## 5 An Application: Dynamic Superimposition of a Knee Joint on a Patient's Leg

The method described for dynamic superimposition is well suited for implementation in complex AR systems. The Virtual Reality Dynamic Anatomy (VRDA) Tool is a system that allows medical practitioners to visualize anatomical structures superimposed on their real counterparts. To realize this effect, the medical practitioner wears a HMD to view a computer graphics model of the knee superimposed on the real leg of a model patient. In the following section, we demonstrate how the method is integrated within the VRDA Tool.

### 5.1 System Setup

The knee is one of the most complex anatomical human joint regarding its structure and its motion. Fortunately, the complexity of motion is not a limiting factor in the proposed method. However, we cannot consider the leg as a rigid object. The muscles and the skin create many perturbations in the 3D marker locations with respect to the bones that we must take into consideration. As a solution to these issues, we treat the leg as two separate objects; the first object is associated with the thigh (femur) and the second object is associated with the shank (tibia and fibula). The 3D models we are using to represent the complete bony knee joint anatomy are high-resolution models from Viewpoint Corporation, acquired by digitizing the anatomy of a cadaver. We employ an OPTOTRAK 3020 optical tracking system that uses active, infrared LEDs as markers. The choice of this system is based upon its resolution, robustness against common perturbations, and speed. The display device is a prototype see-through bench mounted display. We are currently using a Silicon Graphics Deskside Onyx2 with an Infinite Reality2 graphics pipeline to run the application. We perform both computations and stereoscopic rendering on this computer.

### 5.2 Application of the Method and Results

The thigh and shank are tracked independently. To find the best location of the markers, we considered the shape of the leg and selected marker locations where they would probably move least [9]. Also, the correspondence between the real object and the synthetic object is realized by defining common landmarks between the two. We defined these landmarks in places where there is less flesh, allowing the landmarks to be closer to bone to reduce possible scaling or location errors. The synthetic model is scaled based on distance measures between corresponding landmarks on the real leg [26]. The landmarks selected are shown in Figure 1. To determine the relative motion of the markers on the leg, we made 1000 measurements of the global 3D location of the markers over a 10 second interval of standard motion for the leg. We found that the maximum standard deviation of the motion of markers is less than 15 mm.

For the eyepoints, the location from which the projection of the scene is rendered, we chose the center of rotation of the eye [19]. The field of view of the HMD is 26.11° and the display resolution is 640 x 480 pixels. We also applied a coating to the LCD displays to minimize the pixelization of our synthetic objects [18].

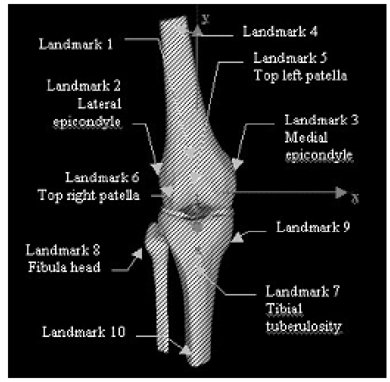


**Figure 1. Ten selected landmarks to associate the real knee with the virtual knee**

**Figure 2. A flexion (left) and extension of the leg of a model patient**

To apply smooth, realistic motion to the knee flexion/extension, we used two lookup tables to determine the locations of the synthetic thigh and shank [6][25]. The lookup tables have 123 entries, corresponding to a 123° range of knee motion. Each entry has six double-precision components, corresponding to rotation and translation about the global x, y, and z axes.

The complete implementation of this method allows superimposition at interactive-speed. We are currently able to achieve frame rates of up to 26.6 Hz, including lag. Furthermore, because of the choice of the SVD method and the enhancement of noise attenuation, the superimposition process is robust and accurate. Two views of the dynamic superimposition are shown in Figure 2.

## 6 Observations and Future Work

In this paper we have presented a robust method that allows real-time, optical superimposition of synthetic objects on dynamic rigid and simple-deformable real objects. Furthermore, we have illustrated these methods with the VRDA Tool, a medical education application for the visualization of internal anatomy on real human anatomy. In the demonstration, we represent the internal motion of the bones of a subject.

**Figure 3. A superimposition with full internal anatomy**

Future work will demonstrate deformable structures such as ligaments and muscles with respect to the bones as well. However, such demonstration is not required in quantifying the methods presented here. Furthermore, methods of non-uniform scaling of synthetic objects will be implemented in future developments. This is especially important in working with generic models that must be registered with specific real objects. Applications of augmented reality methods presented here will be further extended to perform full body motion capture.

## 7 Acknowledgments

The authors thank Ami Sun for providing comprehensive video footage of the VRDA Tool. The work is supported by the National Institute of Health under grant 1-R29-LM06322-01A1, the National Science Foundation under grant EIA-99-86051, the Florida Education Fund, and the Lockheed-Martin Aeronautics Company. Video tapes may be requested at jannick@odalab.ucf.edu .

## References

[1] S. Adelson and L. Hodges. Generating Exact Ray-Traced Animation Frames by Reproduction. *IEEE Computer Graphics and Applications*, 15(3):43–52, 1995.

[2] T. Andriacchi, G. Alexander, M. Toney, C. Dyrby, and J. Sum. A Point Cluster Method for In Vivo Motion Analysis: Applied to a Study of Knee Kinematics. *Journal of Biomechanics*, 1999.

[3] Y. Argotti, V. Outters, L. Davis, A. Sun, and J. Rolland. Technologies for Augmented Reality: Calibration for Real-Time Superimposition on Rigid and Simple-Deformable Real Objects. In *The Fourth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '01), Utrecht, The Netherlands*. Springer-Verlag, October 2001.

[4] Y. Argotti, V. Outters, L. Davis, A. Sun, and J. Rolland. Technologies for Augmented Reality: Real-time Superimposition of Synthetic Objects on Dynamic Rigid and Simple-deformable Real Objects. Technical Report TR01-004, University of Central Florida, 2001.

[5] K. Arun, T. Huang, and S. Blostein. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(5):698–700, 1987.

[6] Y. Baillot, J. Rolland, K. Lin, and D. Wright. Automatic Modeling of Knee-Joint Motion for the Virtual Reality Dynamic Anatomy (VRDA) Tool. *Presence: Teleoperators and Virtual Environments*, 9(3):223–235, 2000.

[7] M. Bajura and U. Neumann. Dynamic Registration Correction in Video-Based Augmented Reality Systems. *IEEE Computer Graphics and Applications*, 15(9):52–60, 1995.

[8] M. Billinghurst and H. Kato. Collaborative Mixed Reality. In *The First International Symposium on Mixed Reality*, pages 261–284, 1999.

[9] A. Cappozzo, A. Cappello, U. Della Groce, and F. Pensalfini. Surface-Marker Cluster Design Criteria for 3-D Bone Movement Reconstruction. *IEEE Transactions on Biomedical Engineering*, 44:1165–1174, 1997.

[10] A. Cappozzo, F. Catani, and A. Leardini. Skin Movement Artifacts in Human Movement Photogrammetry. In *Proceedings of the XIV$^{th}$ Congress of the International Society of Biomechanics*, volume 1, pages 238–239, 1993.

[11] D. Comaniciu, V. Ramesh, and P. Meer. Real-Time Tracking of Non-Rigid Objects Using Mean Shift. In *Proceedings of IEEE Computer Vision and Pattern Recognition*, volume 2, pages 142–149, 2000.

[12] H. Fuchs, M. Livingston, R. Raskar, D. Colucci, K. Keller, A. State, J. Crawford, P. Rademacher, S. Drake, and A. Meyer. Augmented Reality Visualization for Laparoscopic Surgery. In *Proceedings of MICCAI '98*, pages 934–943. Springer-Verlag, 1998.

[13] W. Grimson, R. Kikinis, F. Jolesz, and P. Black. Image Guided Surgery. *Scientific American*, 282(1):63–69, 1999.

[14] G. Halvey and D. Weinshall. Motion of Disturbances: Detection and Tracking of Multi-Body, Non-Rigid Motion. *Machine Vision and Applications*, 11:122–137, 1999.

[15] J. Mellor. Realtime Cameral Calibration for Enhanced Reality Visualization. In *Proceedings of the First International Conference on Computer Vision*. Virtual Reality and Robotics in Medicine, 1995.

[16] W. Robinett and J. Rolland. A Computational Model for Stereoscopic Optics of a Head-Mounted Display. *Presence: Teleoperators and Virtual Environments*, 1(1):45–62, 1992.

[17] J. Rolland, W. Gibson, and D. Ariely. Towards Quantifying Depth and Size Perception in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 4(1):24–49, 1995.

[18] J. Rolland, A. Quinn, K. Arthur, and E. Rinalducci. Quantifying Acuracy and Precision of Rendered Depth in Head-Mounted Displays. Technical Report TR01-002, University of Central Florida, 2001.

[19] J. Rolland and L. Vaissie. Albertian Errors in Head-Mounted Displays: Choice of Eyepoint Location. Technical Report TR01-001, University of Central Florida, 2001.

[20] C. Spoor and F. Veldpaus. Technical Note: Rigid Body Motion Calculated from Spatial Coordinates of Markers. *Journal of Biomechanics*, 13:391–393, 1980.

[21] T. Starner, S. Mann, B. Rhodes, J. Levine, J. Healey, D. Kirsch, R. Picard, and A. Pentland. Augmented Reality Through Wearable Computing. *Presence: Teleoperators and Virtual Environments*, 6(4):386–398, 1997.

[22] A. State, G. Hirota, D. Chen, W. Garrett, and M. Livingston. Superior Augmented-Reality Registration by Integrating Landmark Tracking and Magnetic Tracking. In *Proceedings of SIGGRAPH '96*, pages 429–438. ACM SIGGRAPH, August 1996.

[23] G. Stewart. *Introduction to Matrix Computations*. Academic Press, Inc, New York, 1st edition, 1973.

[24] M. Uenohara and T. Kanade. Vision-Based Object Registration for Real-Time Image Overlay. *International Journal of Computers in Biology and Medicine*, 25(2):249–260, 1995.

[25] S. Van Sint Jan, G. Clapworthy, and M. Rooze. Visualization of Combined Motions in Human Joints. *IEEE Computer Graphics and Applications*, 18(6):10–14, 1998.

[26] D. Wright, J. Renner, J. Jordan, and J. Rolland. Predicting Bone Diameter from Tissue Diameter. *Radiologic Technology*, 2001. In Press.

[27] D. Wright, J. Rolland, and A. Kancherla. Using Virtual Reality to Teach Radiographic Positioning. *Radiologic Technology*, 66(4):167–172, 1995.