

# Non-Rigid Modeling of Body Segments for Improved Skeletal Motion Estimation

Eugene J. Alexander,<sup>1</sup> Christoph Bregler<sup>2</sup> and Thomas P. Andriacchi<sup>3</sup>

**Abstract:** A necessary requirement for many musculoskeletal modeling tasks is an estimation of skeletal motion from observations of the surface of a body segment. The skeletal motion may be used directly for inverse kinematic calculations or as an observation sequence for forward dynamic simulations. This paper describes a fundamentally new approach to human motion capture for biomechanical analysis. Techniques for generating three-dimensional models of human skeletal elements from magnetic resonance imaging data are described, along with a methodology for corresponding these high-resolution internal models to externally observable features. A system for generating dynamic visualizations of these skeletal models from retro-reflective, skin-mounted marker motion capture data is also developed. Next, a set of techniques for estimating body segment shape and pose without the need for retro-reflective markers, from single and multiple, calibrated and un-calibrated cameras is developed. Example results from both synthetic and actual data sequences are presented.

## 1 Introduction

The purpose of this paper is to describe two techniques for human motion capture. The goal of these techniques is to estimate underlying skeletal motion from observations of the surface of the body segment. Both of these techniques model the segments as deformable bodies, improving the accuracy of skeletal motion estimates over existing rigid body modeling approaches. Motion of the body segments does not provide the information needed for biomechanical analysis; what is truly desired is skeletal motion. Estimating skeletal motion from observations

of the segment surface requires sophisticated techniques for modeling the motion of the skin relative to the underlying bone.

Currently one of the primary technical factors limiting the advancement of the study of human movement is the measurement of skeletal movement from markers placed on the skin. The most frequently used method for measuring human movement involves placing retro-reflective markers or fixtures on the skin surface of the segment being analyzed. The movement of the markers is typically used to infer the underlying relative movement between two adjacent segments (e.g. knee joint) with the goal of precisely defining the movement of the joint. Skin movement relative to the underlying skeleton is a primary factor limiting the resolution of detailed joint movement using skin-based systems (Reinschmidt 1997, Holden 1997).

The majority of studies describing three dimensional *in vivo* segment motion do not account for errors associated with non-rigid body movement (Kadaba 1990). Several investigators (Spoor 1980; Cappello 1997) have described methods that were designed to reduce errors associated with non-rigid segment movement. These methods employ skin based marker sets placed on a limb segment to optimally estimate the location and orientation of underlying unobservable skeleton. These techniques in general model the limb segment as a rigid body, then apply various estimation algorithms to obtain an optimal estimate of underlying skeletal motion, subject to a rigid body constraint.

One such rigid body model formulation is given by Spoor and Veldpaus (Spoor 1980); they have described a technique using a minimum mean square error approach that lessens the effect of deformation between any two time steps. This method assumes that deviations from rigidity can be modeled solely by an additive noise. This assumption limits the scope of application for this method, since markers placed directly on skin will actually experience systematic as well as random movement.

---

<sup>1</sup> Div. of Biomechanical Engineering  
Dept. of Mechanical Engineering  
Mail Code: 3030  
Stanford University, Stanford, CA 94301  
Ph: 650-723-9317 Fx: 650-725-1587  
gene.alexander@stanford.edu

<sup>2</sup> Dept. of Computer Science

<sup>3</sup> Dept. of Functional Restoration

Lu expands on the rigid body model approach; rather than seeking optimal rigid body transformations on each segment individually, multiple, constrained rigid body transformations are sought, modeling the hip, knee, and ankle as ball and socket joints (Lu 1999). The assumption of a ball and socket joint is a limitation of this approach for studies of knee motion. Lucchetti takes a different approach, using artifact assessment movements to determine the correlation between flexion–extension angles and skin marker artifact trajectories (Lucchetti 1998). Application of this technique produced decreases in joint center position errors from 14 mm to 4 mm, while orientation errors were reduced from 6 degrees to 3 degrees. A limitation of this approach is the assumption that the skin motion during the dynamic artifact assessment movements is the same as during the dynamic activities. Energy and momentum conserving algorithms for multi-rigid body dynamics have also been developed (Rochinha 2000).

With skin-based marker systems in most cases only large motions such as flexion-extension have acceptable error limits. Cappozzo examined five subjects with external fixator devices and compared the estimates of bone location and orientation between coordinate systems embedded in the bone and coordinate systems determined from skin based marker systems for walking, cycling and flexion-extension activities (Cappozzo 1994). Comparisons of bone orientation from true bone embedded markers versus clusters of three skin-based markers indicate a worst-case root mean square artifact of 7 degrees. They also examined the motion of a number of markers relative to the bone embedded system for markers near bony landmarks and for markers away from such landmarks. A maximum displacement of 40 mm is reported for a marker over the greater trochanter, with markers away from bony landmarks moving appreciably less, with a peak error of approximately 25 mm. Comparisons of bone orientation from true bone embedded markers versus clusters of three skin-based markers indicate a worst-case root mean square artifact of 7 degrees.

Andriacchi described a point cluster technique (PCT) that employs an overabundance of markers (a cluster) placed on each segment to minimize the effects of skin motion artifact (Andriacchi 1994; Andriacchi 1998). The basic PCT can be extended to minimize skin motion artifact by optimal weighting of the markers according to their degree of deformation. Alexander extended the PCT by imposing a functional form on marker motion

relative to the underlying bone (Alexander 2001). By accounting for the changing shape of the limb segment, a decrease in the bone pose estimate error can be achieved. The base PCT transformation equations do not explicitly account for segment deformation; however, the equations can be extended to the general deformation case. Techniques for modeling and subsequently estimating these deformations are summarized in Section 2.3.

Skeletal movement can be measured using alternative approaches to a skin-based marker system. These approaches include stereo radiography (Jonsson 1994), bone pins (LaFortune 1992) external fixation devices (Holden 1997) or single plane fluoroscopic techniques (Banks 1996, Stiehl 1995). While these methods provide direct measurement of skeletal movement, they are invasive or expose the test subject to radiation. These methods also impede natural patterns of movement and care must be taken when attempting to extrapolate these types of measurements to natural patterns of locomotion.

After obtaining an estimate of skeletal movement, it can be beneficial to generate subject-specific geometric representations of anatomy. The development of 3D skeletal models requires reconstruction of the bones and of the articular cartilage. Bone reconstruction is a relatively easy problem, with segmentation possible using standard threshold based, seed growing techniques (Bezdek 93; Warfield 1998; Stammerberger 1999) or more sophisticated model based techniques (Kapur 98, Warfield 2000). Soft tissue structures and particularly articular cartilage models are more difficult to generate accurately. Recent work by Peterfy (Peterfy 1994) exemplifies the new MRI scan sequences and cartilage model generation techniques.

Further developments are needed involving the integration of morphological and material property data with kinematic and kinetic data obtained in a motion laboratory. New developments in medical imaging have made it possible to obtain three-dimensional segmented images of bone, cartilage and connective tissue surrounding joints. The next major developments in this field will likely incorporate information from medical imaging into a construct suitable for functional testing, along with the development of technologies that capture human motion in a completely unrestricted environment.

One such technology is the adaptation of computer vision techniques to biomechanical human motion capture. The vast majority of these modeling techniques are focused

on tracking rigid objects or scenes; however, recent developments in the field hold great promise for tracking deformable objects such as human limbs, faces, torsos, or shod feet. Most techniques treat low-level tracking and 3D structural constraints independently. This paper describes techniques to track and reconstruct non-rigid motions from single or multiple views without prior models. Many non-rigid tracking schemas have been proposed previously. Most techniques use *a priori* models (Lanitis 1995, Blake 1995, DeCarl 1998, Black 1997). Most of these techniques model 2D non-rigid motion, but some of these techniques also recover 3D pose and deformations based on a 3D model. The 3D model is either obtained from 3D scanning devices (Banz 1999) or stereo cameras (Burak 2001), or multi-view reconstruction (Pighin1998, Guenter1998 ). Extensions have been proposed, such as the method for factoring facial expressions and pose during tracking (Bascle 1998). Although it exploits the bi-linearity of 3D pose and non-rigid object configuration, it requires a set of basis images selected before factorization is performed. The discovery of these basis images is not part of the algorithm.

This paper presents a new approach to motion capture that provides a complete solution to the dynamic skeletal modeling problem. Section 2, Development, covers the development of the skeletal models from MRI data, correspondence of the internal model to externally observable features, and the acquisition of motion capture data without the use of retro-reflective markers. Section 3, Results, provide example results for each of these techniques, for both synthetic and actual data. Finally, Section 4, Conclusion, puts these developments in context, describes some limitations of the approach, and proposes some future development directions.

## 2 Development

### 2.1 Skeletal model Generation from MRI Data

Development of three-dimensional subject-specific skeletal models is performed through the use of a model based technique. This technique starts with the construction of a template of the joint from healthy normal data sets. The templates of the knee bones and cartilages were generated by manual segmentation of the MR scan of a normal subject. To facilitate segmentation of cartilage structures from MRI for further image processing, a segmentation technique that is based on

the live wire algorithm (Falcao 1998) was applied.

The inner cartilage surface (ICS) is determined by segmentation and 3D reconstruction of the femur, tibia, and patella. A standard seed growing threshold based segmentation technique is applied to determine the bone voxels. The 3D bone model is then constructed using the marching cubes algorithm (Lorenson 1987).

Initial rigid registration of the bone templates and the reconstructed bone elements of the subject data are accomplished through a 3D surface-based registration technique that does not require the use of an external frame or fiducial markers (Hemler 1995). This detector produces accurate surface points and their corresponding surface normals. This routine is applied to the template and the subject data set. For the model, both the surface points and surface normals are used to form locally supporting planes (for each voxel). These planes form an approximated surface for the baseline skeletal site. For the subject data set, the surface points are matched in the registration procedure onto the surface of the baseline data set. Once these surfaces are detected, we use the Levenberg-Marquardt procedure to find the rigid-body transformation that best matches these two surfaces (Figure 1). The algorithm also produces an outward pointing surface normal at each of the surface points.

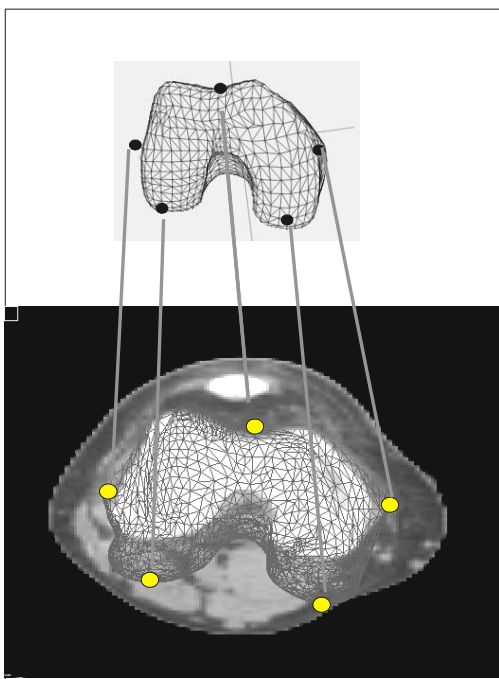
The second stage of the fitting of the bone elements requires an elastic registration approach. We adapted the elastic registration technique described by Dengler (Dengler 1987). This algorithm seeks to find a deformation field  $d(u)$  that minimizes

$$M(u - d(u)) - S(u) \quad (1)$$

where  $M(u)$  is the model data and  $S(u)$  is the subject data. Details are available in (Dengler 1987).

Once the optimal deformation fields have been calculated for the bony elements, they can then be applied to the outward pointing surface normals of the bones and the cartilage templates. The outer cartilage surface (OCS) of the models is then taken as the highest probability location of the subject OCS. A gaussian probability distribution function (GPDF) is centered on this location, directed along the deformed surface normal of the ICS. The variance of the GPDF is selected such that the thickness of the model cartilage is equal to 3 times the standard deviation.

All of the elements required for the final step in the segmentation are now available; the gray scale values of the



**Figure 1** : Correspondence of a three dimensional femoral model to subject MRI data to facilitate cartilage segmentation.

subject data set and a GPDF centered on the initial estimate of the OCS. The Canny edge detection algorithm was modified to incorporate the additional information provided by the GPDF. The gradient of the image data is weighted by the GPDF to provide the final estimate of the OCS. The segmented subject cartilage can now be reconstructed, again using the marching cubes algorithm.

## 2.2 Internal to External Correspondence

The next step is the correspondence of the high-resolution 3D internal images derived from MR scans to externally observable features. The motion of the subject specific 3D anatomic elements can then be driven by data acquired from the motion lab. Fidelity of these animations requires calculation and application of a sequence of rigid body transformations, some of which are directly calculable and some of which are the result of optimizations (the correction for skin marker deformation from rigidity does not use the rigid body assumption, but generates a correction that is applied as a rigid body transform).

In summary, the process is:

- a) Acquire data from MRI and motion capture protocols.
- b) Directly calculate a set of transformations from this data.
- c) Calculate a set of transformations from optimizations.
- d) Generate a 3D representation of the anatomic elements from the MRI data.
- e) Apply the transformation sequence to the geometric representation.

Multiple video cameras acquire data with the subject standing still and during activities of interest. In order to correspond activities in the gait lab with the MR scans, a reference data set (subject standing still, prescribed posture) is acquired with 8 markers clustered about the knee. These markers are filled with a gadolinium solution and covered with a retro-reflective material, rendering them opaque to both imaging modes. The locations of these markers are recorded on the subject with a marker and the circular mounting tapes prior to the subject's transition from the gait lab to the MR lab. The markers are then remounted prior to the MR scan(s).

Each marker is assigned a unit mass and the center of mass and principal axes of inertia are calculated. By treating the center of mass and principal axes as a transformation, local coordinates can be calculated. Another set of coordinate systems is also required for this technique; limb segment specific anatomic landmarks are identified through palpation and a clinically relevant coordinate system defined. The transformations calculated from optimizations bring the anatomic model from the MRI coordinate system to the optimal estimate of the bone location in the global video coordinate system. For the first optimization, for each limb segment, calculate the linear least square error rigid body transformation from the MR limb segment anatomic coordinate system to the video limb segment anatomic coordinate system, using the subset of common markers appropriate for each segment.

The second optimization calculates a correction for the deviation of the limb segment from rigidity during each time step of the activity, using the PCT with either the mass redistribution (Andriacchi 1998) or interval deformation algorithms (Alexander 2001). The interval deformation correction technique used for the experiments in

this paper is summarized section 2.3. The transformation sequence is provided in Table 1. This transformation sequence is applied to each of the anatomic elements over each time step of the activity.

### 2.3 Deformation Correction

A cluster of  $M$  markers is placed on a subject limb segment. The position vector of each marker in the laboratory coordinate system is denoted as  $G(i, n)$  for marker  $i$ , ( $i = 1, 2, \dots, M$ ) at time  $n$ , ( $n = 1, 2, \dots, N$ ), where  $N$  is the number of time samples. A unit weight factor is assigned to each marker for the purpose of calculating the center of mass and principal axes of inertia of the cluster of markers. The global coordinates of each of the markers are given by:

$$G(i, n) = C(n) + E(n)L(i, n) \quad i = 1 \dots M \quad (2)$$

where  $G(i, n)$  is the position vector of marker  $i$ ,  $C(n)$  is the center of mass of the cluster of markers,  $E(n)$  is a rotation matrix constructed from the principal axes of inertia of the cluster, and  $L(i, n)$  is the local position vectors of marker  $i$  in the cluster system.

The measured marker locations  $G(i, n)$  are also defined with respect to the unobservable origin  $P(n)$  and orientation  $O(n)$  of a coordinate system embedded in the bone by

$$G(i, n) = P(n) + O(n)R(i, n) \quad i = 1 \dots M \quad (3)$$

where  $R(i, n)$  is the position vector of marker  $i$  relative to the bone embedded coordinate system at time  $n$ . This bone embedded coordinate system is defined as being coincident with the cluster coordinate system when the subject is at rest in the reference posture. The bone and cluster systems are each orthogonal systems, related by a rigid body transformation which is defined by the rotation  $E_{cb}(n)$  and translation  $C_{cb}(n)$

$$R(i, n) = C_{cb}(n) + E_{cb}(n)L(i, n) \quad i = 1 \dots M \quad (4)$$

Equating 1 and 2, substituting for  $R(i, n)$ , collecting on  $L(i, n)$  and equating like terms yields

$$O(n) = E(n)E_{cb}^{-1}(n) \quad (5a)$$

$$P(n) = C(n) - O(n)C_{cb}(n) \quad (5b)$$

Since  $G(i, n)$  are known by measurement, once  $R(i, n)$  are specified,  $E_{cb}(n)$  and  $C_{cb}(n)$  can be calculated, and subsequently  $O(n)$  and  $P(n)$  are calculable. In other words, once the positions of the markers relative to the bone are successfully estimated, the bone location in the global coordinate system is specified. The remaining task is to determine the motion of the markers relative to the underlying bone. For this problem, the observations are the local position vectors in the cluster system  $L(i, n)$  for all markers  $i$ , all time samples  $N$ . It is assumed that each coordinate  $j$  of each marker  $i$  has some parameterized trajectory  $d(\mathbf{a}_{i,j})$  relative to the underlying bone, with additive noise  $v(i, j, n)$ . These noise distributions are assumed to be independent of each other and are all Gaussian distributions about the parameterized trajectory. Then

$$\hat{R}(i, j, n) = d(\mathbf{a}_{i,j}) + v(i, j, n) \quad j = 1 \dots 3 \quad i = 1 \dots M \quad (6)$$

where  $\mathbf{a}_{i,j}$  is a vector of parameters for marker  $i$ , coordinate  $j$ . Then the estimate of the data is given by

$$\hat{L}(i, n) = E_{cb}^{-1}(n)[\hat{R}(i, n) - C_{cb}(n)] \quad i = 1 \dots M \quad (7)$$

Without further restrictions the problem is indeterminate, as the locations of the markers in the bone system  $R(i, n)$  are never observable with the opto-electronic system. The problem is converted to a chi-squared estimate problem by imposing a functional form on the marker trajectories relative to the bone system and inferring that the observation of the cluster system in the reference position is also an observation of the bone system. The functional form can be selected on the basis of *a priori* knowledge of the activity being studied; for example, in level walking, a periodic activity, the marker motion relative to the underlying bone can be modeled as a sinusoid. The maximum likelihood estimate is then obtained by seeking a minimum to the chi-square error metric:

$$\chi^2 = \sum_{i=1}^M \sum_{j=1}^3 \sum_{n=1}^N \left( \frac{(L(i, j, n) - \hat{L}(i, j, n))}{\sigma(i, j, n)} \right)^2 \quad (8)$$

### 2.4 Markerless Motion Capture – Single Uncalibrated Camera

Given a sequence of  $N$  video frames, the optical flow of  $P$  pixels or the tracked location of  $P$  sparse features in this sequence can be coded as two  $N \times P$  matrices:  $U$  and

From System	To System	Type	Purpose
MR Global	MR Anatomic	Direct	Anatomic element in local system
MR Anatomic	MoCap Anatomic	Optimization	Correspond - internal to external
MoCap Anatomic	MoCap Deformed(t)	Direct	Correspond - Cluster to Bone
MoCap Deformed(t)	MoCap Bone(t)	Optimization	Skin motion correction
MoCap Bone(t)	MoCap Global(t)	Direct	Bone motion

**Table 1** : Transformation sequence. The sequence of transformations required to generate dynamic visualizations of the MRI generated anatomic elements from motion capture data acquired with external retro-reflective markers.

$V$ . Each row of  $U$  holds all x-displacements of all  $P$  locations for a specific time frame, while each row of  $V$  holds all y-displacements for a specific time frame. It has been shown that if  $U$  and  $V$  describe a rigid 3D motion, the rank  $r$  of the composite matrix  $W$ , which is composed of  $U$  stacked on  $V$ , has an upper bound depending on the assumed camera model. For example, for an orthographic camera model the rank  $r \leq 4$ , for a perspective camera model the rank  $r \leq 8$  (Tomasi 1992; Irani 1999). This rank constraint is due to the fact that  $W$  can be factored into two matrices:  $Q \times S \cdot Q$ , which is  $2N \times r$ , describes the relative pose between camera and object for each time frame, and  $S$ , which is  $r \times P$ , describes the 3D structure of the scene that is invariant to camera or object motion.

Bregler has shown that non-rigid object motion can also be factored into two matrices but of rank  $r$  that is higher than the rigid bounds (Bregler 2000). Assuming the 3D non-rigid motion can be approximated by a set of  $K$  modes of variation, the 3D shape of a specific object configuration can be expressed as a linear combination of  $K$  key-shapes ( $S_1, S_2, \dots, S_k$ ). Each key-shape  $S_i$  is a  $3 \times P$  matrix describing  $P$  points. The shape of a specific configuration is a linear combination of this basis set:

$$S = \sum_{i=1}^K l_i \cdot S_i, S_i \in \mathfrak{R}^{3 \times P}, l_i \in \mathfrak{R} \quad (9)$$

Assuming weak-perspective projection, at a specific time frame  $t$  the  $P$  points of a configuration  $S$  project onto 2D image points ( $u_{t,i}, v_{t,i}$ ):

$$\begin{bmatrix} u_{t,1} & \dots & u_{t,p} \\ v_{t,1} & \dots & v_{t,p} \end{bmatrix} = R_t \cdot \left( \sum_{i=1}^K l_{t,i} \cdot S_i \right) + T_t R_t = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \end{bmatrix} \quad (10)$$

$R_t$  contains the first 2 rows of the full 3D camera rotation matrix, and  $T_t$  is the camera translation. The weak

perspective scaling  $f/Z_{avg}$  of the projection is implicitly coded in  $l_{t,1}, \dots, l_{t,K}$ . As in Tomasi-Kanade,  $T_t$  is eliminated by subtracting the mean of all 2D points, and henceforth assume that  $S$  is centered at the origin. Weak perspective projection is in practice a good approximation if the perspective effects between the closest and furthest point on the object surface are small. Extending this framework to full-perspective projection is straightforward using an iterative extension, but all experiments reported here assume weak perspective projection.

We can rewrite this equation as a matrix-matrix multiplication:

$$\begin{bmatrix} u_{t,1} & \dots & u_{t,p} \\ v_{t,1} & \dots & v_{t,p} \end{bmatrix} = \begin{bmatrix} l_{t,1} \cdot R_t & \dots & l_{t,k} \cdot R_t \end{bmatrix} \cdot \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_k \end{bmatrix} \quad (11)$$

By stacking all point tracks from time frame 1 through  $N$  into one large  $2F \times P$  measurement matrix  $W$ , then using equation 11 yields:

$$W = \begin{bmatrix} l_{1,1}R_1 & \dots & l_{1,K}R_1 \\ l_{2,1}R_2 & \dots & l_{2,K}R_2 \\ \vdots & \vdots & \vdots \\ l_{N,1}R_N & \dots & l_{N,K}R_N \end{bmatrix} \cdot \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_K \end{bmatrix} = Q \cdot B \quad (12)$$

Since  $Q$  is a  $2F \times 3K$  matrix and  $B$  is a  $3K \times P$  matrix,  $W$  has, in the noise free case, a rank  $r = 3K$ . The following sections describe how this rank bound on  $W$  can be exploited for the 3D reconstruction of pose, non-rigid deformations, and key-shapes from monocular image sequences. Torresani has previously shown how this constraint can also be used for constrained low-level tracking and recovery of occluded feature locations (Torresani 2001).

The factorization of  $W$  into  $Q$  and  $B$  is not unique. Any invertible  $r \times r$  matrix  $G$  applied as  $(QG^{-1})(GB)$  leads to

an alternative factorization with the same sum-of-squared error as  $QB$ . Using singular value decomposition, we compute a  $\underline{Q}$  with orthonormal columns and a  $\underline{B}$ . In general this  $\underline{Q}$  will not comply to the structure described in equation 12. For the general case, transforming  $\underline{Q}$  into a  $Q$  that complies to these constraints can not be done with a linear least-squares solution. For the non-rigid case, a second factorization step is required on each sub-block that transforms  $\underline{Q}$  into a  $Q$  that complies to the constraints. The details of this second factorization step are described in section 2.6, on multiple calibrated camera homogeneous factorization.

### 2.5 Multiple Un-calibrated Cameras

For biomechanical analysis this factorization technique can be extended to incorporate multi-view inputs from  $M$  cameras. This enlarges the input matrix  $W$  to size  $2FM \times P$ . As in the single view case,  $W$  has  $N$  sub-blocks

$$W_{t,c} = \begin{bmatrix} u_{c,1} & \cdots & u_{c,P} \\ v_{c,1} & \cdots & v_{c,P} \end{bmatrix}, \quad W_t = \begin{bmatrix} W_{t,1} \\ W_{t,2} \\ \vdots \\ W_{t,N} \end{bmatrix}, \quad W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_N \end{bmatrix} \quad (13)$$

As before, we assume that  $W_{t,c}$  can be described by a  $2 \times 3$  pose matrix  $R_{t,c}$ , by  $K$  deformation coefficients  $l_{t,1}, l_{t,2}, \dots, l_{t,k}$ , and a  $3K \times P$  key-shape matrix  $B$ . Assuming the cameras are synchronized, an additional constraint for the multi-view case is that all  $M$  views share the same deformation coefficients for a particular time frame  $t$ :

$$R_t = \begin{bmatrix} R_{t,1} \\ R_{t,2} \\ \vdots \\ R_{t,M} \end{bmatrix}, \quad W_t = [l_{t,1} \cdot R_t \quad l_{t,2} \cdot R_t \quad \cdots \quad l_{t,K} \cdot R_t] \cdot B \quad (14)$$

Similar to our previous 2-step factorization, we can factor  $W$  into  $Q$  and  $B$  complying to this new structure. Furthermore we can enforce another constraint if all  $M$  cameras remain fixed relative to each other: The relative rotation between all  $R_{t,c}$  in the  $R_t$  sub-block of  $Q$  is constant over time. This is enforced with a nonlinear iterative optimization after the 2-step factorization.

### 2.6 Homogeneous factorization from multiple calibrated cameras

With the addition of a calibration procedure it is also possible to extend these techniques to determine the absolute orientation and location relative to some global coordinate system, perhaps one embedded in a force measurement platform. In this case the markers on each limb segment are represented by a set of homogeneous basis shapes  $S_1, S_2, \dots, S_k$ . Each basis shape  $S_i$  is a  $4 \times P$  matrix describing  $P$  homogeneous points. The overall shape is then given by a linear combination of these basis shapes:

$$S, S_i \in \mathbb{R}^{4 \times P}, l_i \in \mathbb{R}S = \sum_{i=1}^K l_i \cdot S_i \quad (15)$$

The  $P$  points of the overall shape  $S$  are projected onto 3D data  $(x,y,z)$ :

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_p \\ y_1 & y_2 & \cdots & y_p \\ z_1 & z_2 & \cdots & z_p \end{bmatrix} = T \cdot S, \quad T = \begin{bmatrix} r_1 & r_2 & r_3 & t_x \\ r_4 & r_5 & r_6 & t_y \\ r_7 & r_8 & r_9 & t_z \end{bmatrix} \quad (16)$$

Equation (16) can be rewritten as a matrix-matrix multiplication

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_p \\ y_1 & y_2 & \cdots & y_p \\ z_1 & z_2 & \cdots & z_p \end{bmatrix} = [l_1 \cdot T \quad \cdots \quad l_k \cdot T] \cdot \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_k \end{bmatrix} \quad (17)$$

Add a time index to each 3D point and denote point  $i$  in frame  $t$  as  $(x_i^{(t)}, y_i^{(t)}, z_i^{(t)})$ . Code the  $N$  frames of data into a matrix and use Eq. 17:

$$\begin{bmatrix} x_1^{(1)} & \cdots & x_p^{(1)} \\ y_1^{(1)} & \cdots & y_p^{(1)} \\ z_1^{(1)} & \cdots & z_p^{(1)} \\ \vdots & \cdots & \vdots \\ x_1^{(N)} & \cdots & x_p^{(N)} \\ y_1^{(N)} & \cdots & y_p^{(N)} \\ z_1^{(N)} & \cdots & z_p^{(N)} \end{bmatrix} = \underbrace{\begin{bmatrix} l_1^{(1)} \cdot T^{(1)} & \cdots & l_k^{(1)} \cdot T^{(1)} \\ \vdots & \cdots & \vdots \\ l_1^{(N)} \cdot T^{(N)} & \cdots & l_k^{(N)} \cdot T^{(N)} \end{bmatrix}}_Q \cdot \underbrace{\begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_k \end{bmatrix}}_B \quad (18)$$

Eq. 18 shows that the tracking matrix can be factored into 2 matrices:  $Q$  contains the pose  $T$  and configuration weights  $l_i$  for each time step, while  $S$  contains the basis shapes. This factorization can be performed by singular value decomposition, only using the first  $4K$  singular vectors:  $W^{(3N \times P)} = U * D * V^T = Q^{(3N \times 4K)} * S^{(4K \times P)}$ .

The second step in the process is to extract the pose  $T$  and configuration weights  $l_i$  from the matrix  $Q$ . This can be done by considering a sub-matrix of  $Q$  corresponding to a single time interval,  $q^{(t)}$ , re-ordering that sub-matrix,  $\underline{q}^{(t)}$ , then again applying the SVD as follows:

$$q^{(t)} = \begin{bmatrix} l_1^{(t)} \cdot T^{(t)} & \dots & l_K^{(t)} \cdot T^{(t)} \end{bmatrix} = \begin{bmatrix} l_1^{(t)} \cdot r_1^t & l_1^{(t)} \cdot r_2^t & l_1^{(t)} \cdot r_3^t & l_1^{(t)} \cdot t_x^t & \dots & l_K^{(t)} \cdot r_1^t & l_K^{(t)} \cdot r_2^t & l_K^{(t)} \cdot r_3^t & l_K^{(t)} \cdot t_x^t \\ l_1^{(t)} \cdot r_4^t & l_1^{(t)} \cdot r_5^t & l_1^{(t)} \cdot r_6^t & l_1^{(t)} \cdot t_y^t & \dots & l_K^{(t)} \cdot r_4^t & l_K^{(t)} \cdot r_5^t & l_K^{(t)} \cdot r_6^t & l_K^{(t)} \cdot t_y^t \\ l_1^{(t)} \cdot r_7^t & l_1^{(t)} \cdot r_8^t & l_1^{(t)} \cdot r_9^t & l_1^{(t)} \cdot t_z^t & \dots & l_K^{(t)} \cdot r_7^t & l_K^{(t)} \cdot r_8^t & l_K^{(t)} \cdot r_9^t & l_K^{(t)} \cdot t_z^t \end{bmatrix} \quad (19)$$

$$\underline{q}^{(t)} = \begin{bmatrix} l_1^{(t)} \cdot r_1^t & l_1^{(t)} \cdot r_2^t & l_1^{(t)} \cdot r_3^t & l_1^{(t)} \cdot t_x^t & l_1^{(t)} \cdot r_4^t & l_1^{(t)} \cdot r_5^t \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ l_K^{(t)} \cdot r_1^t & l_K^{(t)} \cdot r_2^t & l_K^{(t)} \cdot r_3^t & l_K^{(t)} \cdot t_x^t & l_K^{(t)} \cdot r_4^t & l_K^{(t)} \cdot r_5^t \\ & l_1^{(t)} \cdot r_6^t & l_1^{(t)} \cdot t_y^t & l_1^{(t)} \cdot r_7^t & l_1^{(t)} \cdot r_8^t & l_1^{(t)} \cdot r_9^t & l_1^{(t)} \cdot t_z^t \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ & l_K^{(t)} \cdot r_6^t & l_K^{(t)} \cdot t_y^t & l_K^{(t)} \cdot r_7^t & l_K^{(t)} \cdot r_8^t & l_K^{(t)} \cdot r_9^t & l_K^{(t)} \cdot t_z^t \end{bmatrix} \quad (20)$$

$$= \begin{bmatrix} l_1^{(t)} \\ \vdots \\ l_K^{(t)} \end{bmatrix} \cdot \begin{bmatrix} r_1^t & r_2^t & r_3^t & t_x^t & r_4^t & r_5^t & r_6^t & t_y^t & r_7^t & r_8^t & r_9^t & t_z^t \end{bmatrix} \quad (21)$$

This procedure is repeated for all time blocks, yielding the pose  $T^{(t)}$  and configuration weights  $l_1^{(t)} \dots l_K^{(t)}$ . This decomposition is not unique, but by enforcing the known constraints of a homogeneous transformation matrix, a transformation  $G$  that maps  $\underline{T}$  into a homogeneous  $T^{(t)} = \underline{T}^{(t)} * G$  can be determined. This transformation must also be applied to the basis shape,  $S = G^{-1} * \underline{S}$ .

The choice of the number of basis shapes  $K$  allow for a selection of deformation correction capability appropriate to the data. As  $K$  increases the ability of the algorithm to model deformation increases, up to a limit dependant on the number of markers on the limb segment.

### 3 Results

#### 3.1 Retro-reflective marker correspondence to MRI data

Integration of the high resolution MR imaging of the bones and articular cartilage of the knee with retro-reflective marker kinematic measurements was tested.

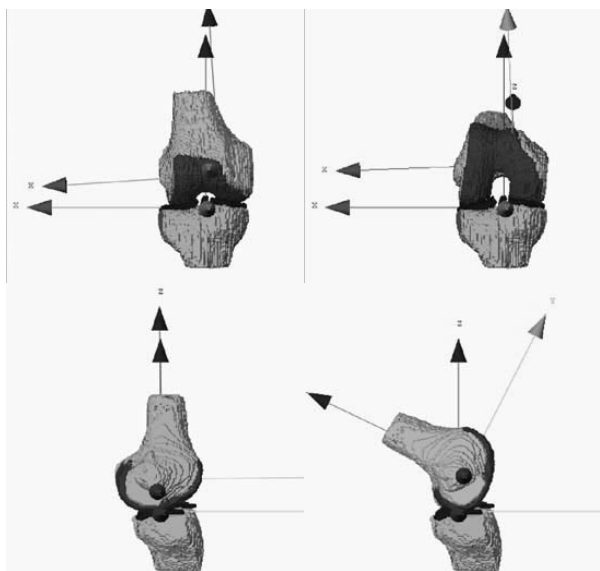
External markers filled with Gd-DTPA (Magnevist(r), Berlex Inc., Wayne, N.J.) doped water were applied to the skin around the knee joint at the same positions used for gait analysis in the Biomotion Laboratory (Figure 2). Three volunteers underwent MRI of the knee joint in this fashion (GE Signa, 1.5T, 3D SPGR, TR=60msec, TE=5msec, flip angle 40°, 1 excitation, matrix 256x160 elements, rectangular FOV 16x12 cm, slice thickness 1.3 mm, 128 slices, with fat saturation and repeated without fat saturation). After image acquisition, the MR images were transferred to an independent imaging workstation (Advantage Windows, General Electric, Madison, WI) and the femoral and tibial bones and cartilage were segmented. 3D reconstructions of the femoral and tibial cartilage and of the femoral and tibial bone were generated in this fashion.



**Figure 2** : Gadolinium filled retro-reflective marker set affixed the knee of a test subject during a static reference pose. The markers are re-mounted prior to acquisition of MRI data but removed from the subject prior to any dynamic activities.

Each subject was tested standing still, during level walking, seated leg extension, standing leg flexion, ascending and descending stairs. The level walking measurements included six stride cycles over a range of walking speeds. The instrumentation included a four-camera, video-based, optoelectronic system for 3-D motion analysis, a multicomponent force plate for measurement of foot-ground reaction force and a computer system for acquisition, processing and analysis. The experimental model used for the functional evaluation study idealizes the lower extremity as 3 segments with six degree-of-freedom joints at the knee and ankle. The data for each of the activities was reduced to six degree of freedom motion of the thigh and shank segment, including application of the interval deformation correction technique to minimize the effects of segment deformation. An exam-

ple of the resultant kinematic functional joint images is demonstrated in Figure 3.



**Figure 3** : Two frames from a visualization of a healthy knee performing a leg extension in a fixed tibial reference frame. The top two figures are from an anterior view; the bottom two, a lateral view. The full movie is available at: [http://biomotion.stanford.edu/CMES\\_2001.htm](http://biomotion.stanford.edu/CMES_2001.htm)

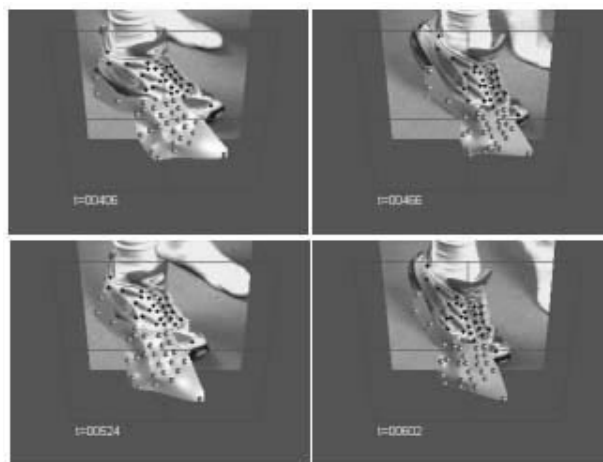
### 3.2 Rank-Constrained Tracking and 3D reconstruction from single uncalibrated camera

The rank-constraint technique for optical flow estimation was tested on a 500 frame long video sequence of a deforming shoe. The recordings are challenging due to changes in the object appearance that are caused by the large rotations and deformations as well as by variations in illumination. In this example a set of 30 reliable features were initially tracked using the technique of Lucas-Kanade and employing affine transformations for the patches centered at these points. The reference patch of each point was updated every 10 frames in order to accommodate the changes in feature appearance of our long sequence. The multi-frame approach was capable of recovering the possible drifting introduced on some of the tracks by the frequent update of the template for the points. 80 additional features were then selected along 1D edges in the reference frame.

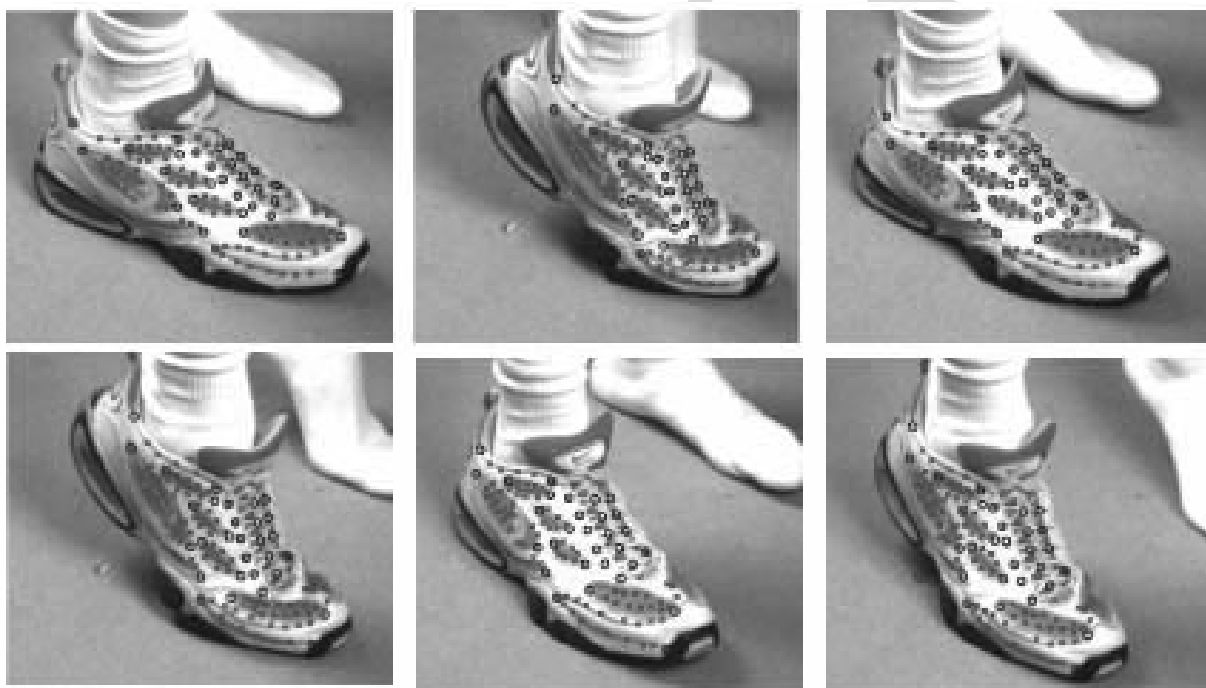
A first approximate initialization of their displacements was produced by linear extrapolation from the motion of

the reliable points. The resulting  $W$  matrix was used as an initialization for the tracking technique based on rank constraints. Experimentation with different values for the rank achieved the best solution by setting it to rank  $r = 9$ . The classic pyramidal approach was employed in smoothing the images and several iterations of the multi-frame method were run. The algorithm robustly and very accurately tracked most of the 110 points throughout the whole sequence. Figure 4 shows the features tracked for several frames.

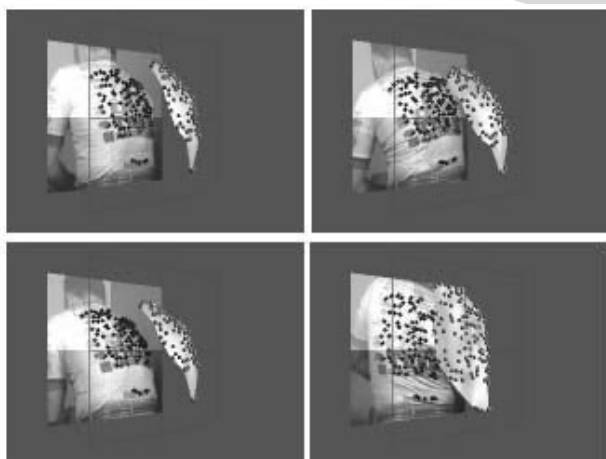
Given the estimated  $Q$  and  $B$  of those 500 tracked monocular image frames, the single-camera 3D reconstruction technique was then applied. Figure 5 shows some example frames with the reconstructed non-rigid 3D shapes overlaid. The new reconstruction technique was then applied to a video recording of a deforming human torso, one of the most challenging body segments to model using retro-reflective marker based systems due to the high degree of non-rigidity. The subject is wearing a tight fitting spandex bicycling jersey, with the features to be tracked selected from elements of the text on the jersey. Figure 6 shows some example frames with the reconstructed non-rigid 3D shapes overlaid.



**Figure 5** : Along with tracking, the rank constrained algorithm also produces an estimate of the deforming shape of the object. This sequence of four images shows the shoe rotating and deforming. The full movie is available at: [http://biomotion.stanford.edu/CMES\\_2001.htm](http://biomotion.stanford.edu/CMES_2001.htm)



**Figure 4 :** Six frames from the tracking of the deforming, rotating shoe sequence. The hollow, blue circles are features with good 2D texture; the filled, red circles are features with 1D texture that the rank constrained algorithm was able to track. The full movie is available at: [http://biomotion.stanford.edu/CMES\\_2001.htm](http://biomotion.stanford.edu/CMES_2001.htm)



**Figure 6 :** The rank constrained algorithm applied to an image sequence of a rotating, deforming torso. The features are selected from points on the text on the jersey. The full movie is available at: [http://biomotion.stanford.edu/CMES\\_2001.htm](http://biomotion.stanford.edu/CMES_2001.htm)

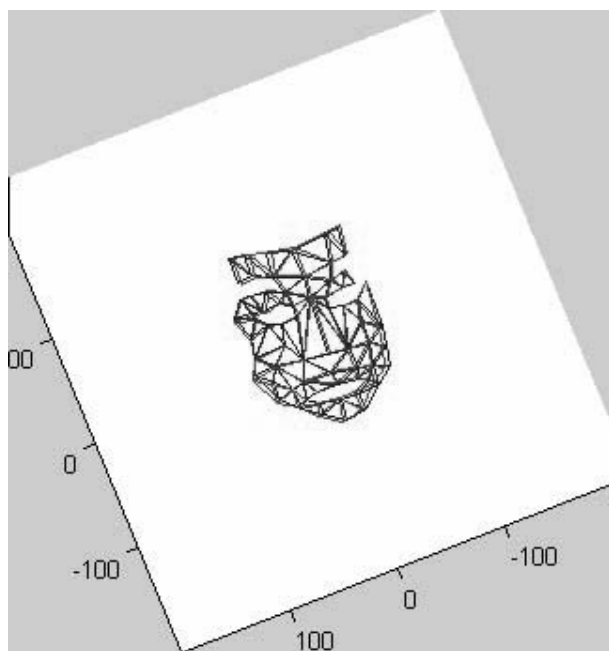
Approach or # Basis Shapes	Spoor & Veldpas	1	2	3	4	5
Error (% of face size)	3.18	3.28	2.95	2.65	2.46	2.26

**Table 2 :** Error chart, rigid body model vs. multiple uncalibrated camera 3D reconstruction, real data. The root mean square 3D reconstruction error per marker per time step for the Spoor & Veldpas rigid body technique compared to the homogeneous factorization technique for the actual data sequence of the face. As the number of basis shapes is increased the error decreases.

### 3.3 3D Reconstruction from Multiple uncalibrated cameras

Table 2 shows the multiple uncalibrated camera reconstruction errors for a video sequence of a face rotating and deforming. This reconstruction is the most challenging task, since it tests the entire system from video input to 3D output. Overall the small reconstruction errors tell us that this technique is indeed able to accurately recover non-rigid deformations from multiple uncalibrated camera image sequences. As shown in Figure 7, both the

overall orientation of the head and the detailed motion of the lips and eyebrows are accurately reconstructed with this technique.



**Figure 7** : The multiple uncalibrated camera algorithm applied to a two camera view of facial rotation and deformation. The red indicates the ground truth data, the blue the reconstructed data with 5 key-shapes. The full movie is available at: [http://biomotion.stanford.edu/CMES\\_2001.htm](http://biomotion.stanford.edu/CMES_2001.htm)

### 3.4 3D Shape, Absolute Orientation and Location from Multiple Calibrated Cameras

20 simulated data sets were generated and examined with the Spoor rigid body approach (Spoor 1980) and the homogeneous factorization approach presented here. The homogeneous factorization approach was tested for five different numbers of basis shapes, from one to five. The data sets were contaminated with systematic and random error to mimic the muscle deformation, skin sliding, skin jiggling error commonly seen in actual motion capture data streams.

The two software correction techniques were compared to the known, generated data. The root mean error in reconstructing the locations of each of the markers for each time step of each of the data sets was calculated for both of the estimation techniques. Table 3 summarizes the results for all of the data sets.

Approach or # Basis Shapes	Spoor & Veldpas	1	2	3	4	5
Error (mm)	6.8	7.2	6.7	6.4	6.3	6.2

**Table 3** : Error chart, rigid body model vs. multiple calibrated camera homogeneous factorization, synthetic data. The root mean square 3D reconstruction error per marker per time step for the Spoor & Veldpas rigid body technique compared to the homogeneous factorization technique for the synthetic data sequence. As the number of basis shapes is increased the error decreases.

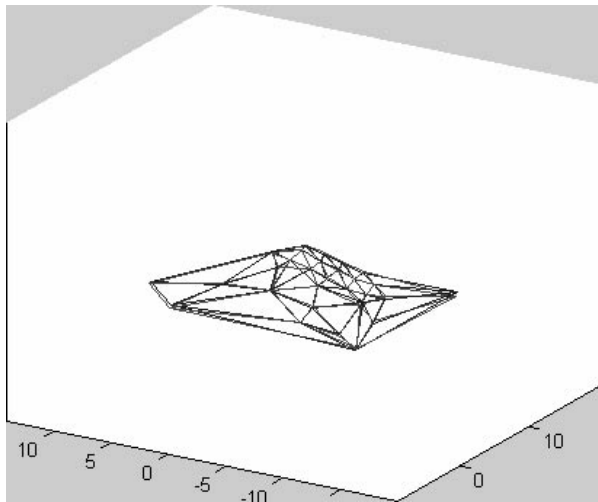
The homogeneous factorization technique with multiple calibrated cameras was also applied to the shoe data sequence. Table 4 summarizes the results, showing the accuracy increase as the number of key shapes is increases. Figure 8 shows a sample video frame, with the ground truth data in red and the reconstruction in blue. For this sequence the best results were obtained with the rank set to  $r = 5$ .

Approach or # Basis Shapes	Spoor & Veldpas	1	2	3	4	5
Error (mm)	2.3	2.6	2.3	2.0	1.8	1.7

**Table 4** : Error chart, rigid body model vs. multiple calibrated camera homogeneous factorization, actual data. The root mean square 3D reconstruction error per marker per time step for the Spoor & Veldpas rigid body technique compared to the homogeneous factorization technique for the shoe data sequence. As the number of basis shapes is increased the error decreases.

## 4 Discussion

Since inception, the field of human motion capture for biomechanical analysis has been dominated by video capture of retro-reflective markers and generic models of bones. For the first time, this paper has demonstrated techniques for generating subject-specific skeletal models from MRI data, corresponding these models to externally observable features on non-rigid body segments, and tracking these deforming segments from single and multiple, calibrated and uncalibrated cameras. The recovery of 3D shape, orientation, and location of these body parts without the use of retro-reflective markers



**Figure 8 :** Application of the homogeneous factorization algorithm to the deforming shoe sequence. The red indicates the ground truth data, the blue the reconstructed data with 5 key-shapes. The full movie is available at: [http://biomotion.stanford.edu/CMES\\_2001.htm](http://biomotion.stanford.edu/CMES_2001.htm)

holds great promise for the future of biomechanical analysis of human motion.

There are a number of limitations to the techniques described here. Fixation of the gadolinium markers to the knee during the MR sequence results in a requirement to acquire data on a larger volume than would be needed otherwise. The result is either longer (in time) scan sequences or a decrease in resolution. In future work we plan to experiment with smaller markers. We have not yet determined a technique for automatically selecting the optimal rank for the factorization. We also do not yet have a technique for automatically corresponding feature points from multiple cameras, fortunately, this manual correspondence only need to be performed on one video frame. Some questions are undoubtedly raised about the appropriateness of tracking features on tight fitting clothing versus the standard technique of observing markers on the skin. One argument in favor of the clothing approach is that the material properties of the observable object can be known exactly, as opposed to the actual skin of the subject. Of course this technique does not actually require clothing; subject skin could be observed directly, with the features to be tracked drawn on the skin using a surgical marker.

Future work will focus on some of the limitations mentioned; automatic correspondence between cameras and

automatic optimal rank determination. Other areas for future work include validation studies on skeletal motion estimation, perhaps on subjects with external bone fixators or data simultaneously acquired with bi-planar radiographic techniques.

Reconstructing non-rigid models of human body segments without the need for retro-reflective markers has many potential applications in biomechanical modeling. The techniques described in this work should provide a solid foundation for future clinical studies in orthopedics and neurology, sports injury prevention, sports performance, and many other areas.

**Acknowledgement:** This work was supported in part by grants from the National Institutes of Health, Defense Advanced Research Projects Agency, and the Stanford BioX Interdisciplinary Initiative Program. The assistance of Danny Yang, Lorenzo Torresani, James Davis, Ajit Chaudhari, Rich Bragg, and Chris Dyrby is gratefully acknowledged.

## References

- Alexander, E.J., Andriacchi, T.P.** (2001): Correcting for Deformation in Skin Based Marker Systems, *Journal of Biomechanics*, 34(3) 355-362.
- Andriacchi, T.P., Sen, K., Toney, M.K. and Yoder, D.** (1994): New Developments in Musculoskeletal Testing, *Proceeding of the Canadian Society of Biomechanics*, 221-222.
- Andriacchi, T.P., Alexander, E.J., Toney, M.K., Dyrby, C.O. and Sum, J.** (1998): A Point Cluster Method for *In vivo* Motion Analysis: Applied to a Study of Knee Kinematics. *Journal of Biomechanical Engineering* 120(12), 743 - 749.
- Banks, S.A.; Hodge, W.A.** (1996): Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy. *IEEE Transactions on Biomedical Engineering* 43(6), 638-649.
- Bascle, B.; Blake, A.** (1998): Separability of Pose and Expression in facial tracking and animation, *Proceedings of the International Conference on Computer Vision 1998*.
- Bezdek, J.C.; Hall, L.O.; Clarke, L.P.** (1993): Review of MR image segmentation techniques using pattern recognition. *Medical Physics*, 20(4):1033 - 1048.

- Black, M.J.; Yacoob, Y.; Jepson, A.D.; Fleet, D.J.** (1997): Learning Parameterized Models of Image Motion, *Proceedings of the Computer Vision and Pattern Recognition Conference 1997*.
- Blake, A.; Isard, M.; Reynard, D.** (1995): Learning to track the visual motion of contours, *Journal of Artificial Intelligence*.
- Blanz, V.; Vetter, T.** (1999): A Morphable Model for the Synthesis of 3D Faces, *Proceedings of SIGGRAPH Conference 1999*.
- Bregler, C.; Hertzmann, A.; Biermann, H** (2000): Recovering Non-Rigid 3D Shape from Image Streams, *Proceedings of the Computer Vision and Pattern Recognition Conference 2000*.
- Cappozzo, A., Cappello, A., Della Croce, U. and Pensalfini, F.** (1997): Surface Marker Cluster Design Criteria for 3-D Bone Movement Reconstruction. *IEEE Transactions on Biomedical Engineering*. 44, 1165-1174.
- Capozzo, A., Catani, F., Leardini, A., Benedetti, M.G. and Della Croce, U.** (1996): Position and Orientation in Space of Bones During Movement: Experimental Artefacts. *Clinical Biomechanics* 11: 90-100.
- Cappello, A., Cappozzo, A., La Palombara, P.F., Lucchetti, L. and Leardini, A.** (1997): Multiple anatomical landmark calibration for optimal bone pose estimation, *Human Movement Science* 16(2-3) 259-274.
- DeCarlo, D.; Metaxas, D.** (1998): Deformable Model-Based Shape and Motion Analysis from Images using Motion Residual Error, *Proceedings of the International Conference on Computer Vision 1998*.
- Dengler J.; Schmidt, M.** (1987): The Dynamic Pyramid – A Model for Motion Analysis with Controlled Continuity, *International Journal of Pattern Recognition and Artificial Intelligence*, 2, pp. 275 – 286.
- Falcão A.X.; Udupa J.K.; Samarasekera S.; Sharma S.** (1998): User steered image segmentation paradigms: Live wire and live lane. *Graphical Models and Image Processing*; 60: 233-260.
- Gokturk, S.B.; Bouget, J.Y.; Grzeszocuk, R.** (2001): A Data-Driven Model for Monocular Face Tracking, *Proceedings of the International Conference on Computer Vision, 2001*.
- Guenter, B.; Grimm, C.; Wood, D.; Malvar, H.; Pighin, F.** (1998): Making Faces, *Proceedings of the SIGGRAPH Conference 1998*.
- Hemler P.F.; Napel S.; Sumanaweera T.S.; Pichumani R.; van den Elsen P.A.; Martin D.,; Drace J.; Adler J.R.; Perkash I.** (1995): Registration error quantification of a surface-based multimodality image fusion system. *Medical Physics* 22 (7): 1049-56.
- Holden, J.P., Orsini, J.A., Siegel, K.L., Kepple, T.M., Gerber, L.H. and Stanhope, S.J.** (1997): Surface movement errors in shank kinematics and knee kinematics during gait *Gait and Posture* 5, 217-227.
- Irani, M.** (1999): Multi-Frame Optical Flow Estimation Using Subspace Constraints, *Proceedings of the International Conference on Computer Vision 1999*.
- Jonsson, H.; Karrholm, J.** (1994): Three-dimensional Knee Joint Movements During a Step-up: Evaluation after Cruciate Ligament Rupture. *Journal of Orthopedic Research* 12(6): 769-779.
- Kadaba, M.P., Ramakrishnan, M.K. and Wooten, M.E.** (1990): Measurement of lower extremity kinematics during level walking, *Journal of Orthopedic Research* 8, 383-392.
- Kapur, T; Beardsley, P.; Gibson, S.; Grimson, W; Wells, W.** (1998): Model-based segmentation of clinical knee MRI. *Proceedings of the IEEE International Workshop on Model-Based 3D Image Analysis*, pages 97—106.
- Lafortune, M.A., Cavanagh, P.R., Sommer, H.J. and Kalenak, A.** (1992): Three dimensional kinematics of the human knee during walking. *Journal of Biomechanics* 25, 347-357.
- Lanitis, A.; Taylor C.J.; Cootes T.F.; Ahmed T.** (1995): Automatic Interpretation of Human Faces and Hand Gestures Using Flexible Models, *International Workshop on Automatic Face- and Gesture-Recognition 1995*.
- Lorensen, W.E.; Cline H.E.** (1987): Marching Cubes: A High Resolution 3D Surface Construction Algorithm, *Proceedings of SIGGRAPH 1987*, Vol. 21, No. 4, pp. 163-169.
- Lu, T.W. and O'Connor, J.J.** (1999): Bone position estimation from skin marker co-ordinates using global optimization with joint constraints. *Journal of Biomechanics* 32, 129 – 134.
- Lucchetti, L., Cappozzo, A., Cappello, A. and Della Croce, U.** (1998): Skin movement artefact assessment

- and compensation in the estimation of knee-joint kinematics. *Journal of Biomechanics* 31, 977-984.
- Peterfy C.G.; Dijke C.F.; van Janzen D.L.; Gluer C.C.; Namba R.; Majumdar S.; Lang P.; Genant H.K.** (1994): Quantification of articular cartilage in the knee with pulsed saturation transfer subtraction and fat-suppressed MR imaging: optimization and validation. *Radiology*; 192:485-491.
- Pighin, F.; Hecker, J.; Lischinski, D.; Szeliski, R.; Salesin, D.** (1998): Synthesizing Realistic Facial Expressions From Photographs, *Proceedings of the SIGGRAPH Conference 1998*.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P.** (1996): *Numerical Recipes in C: The Art of Scientific Computing* (2nd Edn), pp. 444-455. Cambridge University Press, New York.
- Reinschmidt, C., van den Bogert, A.J., Nigg, B.M., Lundberg, A. and Murphy, N.** (1997): Effect of Skin Movement on the Analysis of Skeletal Knee Joint Motion During Running. *Journal of Biomechanics* 30, 729 – 732.
- Rochinha, F.A.; Sampaio, R.** (2000): Nonlinear Rigid Body Dynamics: Energy and Momentum Conserving Algorithm. *CMES: Computer Modeling in Engineering and Sciences*, Vol 1, No 2, pp 7-18.
- Spoor, C.W. and Veldpaus, F.E.** (1980): Rigid Body Motion Calculated from Spatial Coordinates of Markers. *Journal of Biomechanics* 13, 391-393.
- Stammlerger T.; Eckstein F.; Michaelis M.; Englmeier K.H.; Reiser M.** (1999): Interobserver reproducibility of quantitative cartilage measurements: comparison of B-spline snakes and manual segmentation. *Magnetic Resonance Imaging*, 17:1033-1042.
- Stiehl J.B., Komistek R.D., Dennis D.A., Paxson R.D., Hoff W.A.** (1995): Fluoroscopic analysis of kinematics after posterior-cruciate-retaining knee arthroplasty. *Journal of Bone and Joint Surgery* 77, 884-889.
- Tomasi, C.; Kanade, T.** (1992): Shape and motion from image streams under orthography: a factorization, *International Journal of Computer Vision*, 9(2), 137-154.
- Torresani, L.; Yang D.; Alexander, E.J.; Bregler, C.** (2001): Tracking and Modeling Non-Rigid Objects with Rank Constraints, in review, *Proceedings of the Computer Vision and Pattern Recognition Conference*.
- Warfield, S.K.; Kaus, M.; Jolesz, F.A.; Kikinis** (2000): Adaptive, Template Moderated, Spatially Varying Statistical Classification. *Medical Image Analysis* 4(1), pp 43-55.
- Warfield S.; Winalski C.; Joesz F.; Kikinis R.** (1998): Automatic segmentation of MRI of the knee, *Proceedings of the 6th Annual Meeting of ISMRM*, Sydney, Australia, p 563.