Measuring human movement for biomechanical applications using markerless motion capture

Lars Mündermann*[a], Stefano Corazza[a], Ajit M. Chaudhari[a], Thomas P. Andriacchi[a,b,c], Aravind Sundaresan[d], Rama Chellappa[d]

[a]Department of Mechanical Engineering, Stanford University, Stanford, CA
[b]Bone and Joint Center, Palo Alto VA, Palo Alto, CA
[c]Department of Orthopedic Surgery, Stanford University Medical Center, Stanford, CA
[d]Center for Automation Research, University of Maryland, College Park, MD
*lmuender@stanford.edu; biomotion.stanford.edu

ABSTRACT

Modern biomechanical and clinical applications require the accurate capture of normal and pathological human movement without the artifacts associated with standard marker-based motion capture techniques such as soft tissue artifacts and the risk of artificial stimulus of taped-on or strapped-on markers. In this study, the need for new markerless human motion capture methods is discussed in view of biomechanical applications. Three different approaches for estimating human movement from multiple image sequences were explored. The first two approaches tracked a 3D articulated model in 3D representations constructed from the image sequences, while the third approach tracked a 3D articulated model in multiple 2D image planes. The three methods are systematically evaluated and results for real data are presented. The role of choosing appropriate technical equipment and algorithms for accurate markerless motion capture is critical. The implementation of this new methodology offers the promise for simple, time-efficient, and potentially more meaningful assessments of human movement in research and clinical practice.

Keywords: Motion capture, movement analysis, markerless, articulated model, visual hull, iterative closest point, simulated annealing, optical flow

1. INTRODUCTION

Human movement analysis spans many different fields such as kinesiology, physiotherapy, orthopedic surgery, ergonomics, etc [1-7]. Moreover, human motion capture is a well established paradigm for the diagnosis of the patho-mechanics related to musculoskeletal diseases, the development and evaluation of rehabilitative treatments and preventive interventions for musculoskeletal diseases. Over the last several centuries our understanding of human locomotion has been a function of the methods to capture human movement that were available at the time. The Weber brothers (1836) reported one of the first quantitative studies of the temporal and distance parameters during human locomotion [8]. The works of two contemporaries, Marey (1873) and Muybridge (1878), were among the first to quantify patterns of human movement using photographic techniques [9, 10].

In many cases the expanded need for enhancing our understanding of normal and pathological human movement drove the introduction of new methods to capture human movement. At present, the most common methods for accurate capture of three-dimensional human movement require a laboratory environment and the attachment of markers, fixtures or sensors to body segments. The constraints of the laboratory environment as well as the markers placed on the subjects can mask subtle but important changes to the patterns of locomotion. Previous experience has demonstrated that minor changes in patterns of locomotion can have a profound impact on the outcome of treatment or progression of musculoskeletal pathology. For example, it has been shown that the mechanics of walking was changed in patients with anterior cruciate ligament deficiency of the knee [11]; functional loading influenced the outcome of high tibial osteotomy [12]; functional performance of patients with total knee replacement was influenced by the design of the implant [13], and the mechanics of walking influenced the progression of osteoarthritis of the knee [14]. Each of the clinical examples referenced above were associated with subtle but important changes to the mechanics of walking. In general, human movement analysis refers to studies on whole-body human motion with applications to posture studies, human identification, and detection of abnormal gait. For example, precise characterization and diagnosis of movement disorders is possible using stereophotogrammetry which records movement with small adhesive markers placed on body segments. However, widespread use in the medical setting is severely limited by the cumbersome, time-consuming and invasive nature of marker-based systems. Thus, the ability to measure patterns of locomotion without the risk of an artificial stimulus producing unwanted artifacts that could mask the natural patterns of motion is an important need for emerging health care applications. Ideally, the measurement system/protocol should be neither invasive nor harmful and only minimally encumber the subject. Furthermore, it should allow measuring subjects in their natural environment such as their work place, home, or on sport fields and be capable of measuring natural activities/motion over a sufficiently large field of view.

Our understanding of normal and pathological human movement would be enhanced by a method that allows the capture of human movement without the constraint of markers or fixtures placed on the body. To date, markerless methods are not widely available because the accurate capture of human movement without markers is technically challenging yet recent technical developments in computer vision provide the potential for markerless human motion capture for biomechanical and clinical applications. The purpose of this study was to evaluate the accuracy of human body kinematics extracted through a markerless motion capture system in view of biomechanical applications. Several approaches for tracking human body segments for determining human body kinematics are discussed and evaluated. The principal goal of this collaborative research effort is to design, develop and evaluate novel systems using multiple optical sensors that will efficiently and accurately provide 3D measurements of human movement. The availability of such systems offer the promise of expanding the applicability of methods for diagnosis and treatment of movement-related disorders, minimizing patient preparation time, and reducing experimental artifacts caused by, for instance, inter-observer variability. Thus, the implementation of this new technology will allow for simple, time-efficient, and potentially more meaningful assessments of gait in research and clinical practice.

The notions of markerless motion capture have already appeared in computer vision and biomechanical literature. Following their review (Section 2), we outline our approaches (Section 3). In Section 4, we present results. We conclude the paper with a discussion of the results, and problems open for further research (Section 5).

2. BACKGROUND

The development of markerless motion capture systems did not originate from the clinical perspective but within the fields of computer vision and machine learning, where the analysis of human actions by a computer is gaining increasing interest. Potential applications of human motion capture are the driving force of system development, and the major application areas are: smart surveillance, control, perceptual interface, virtual reality, view interpolation, and motion analysis. One of the key challenges in a markerless system is the acquisition and representation of human movement. Vision systems are typically divided into two categories, namely active and passive vision system. Active systems emit light-information in the visible or infrared light spectrum, while passive systems rely purely on capturing images. The main focus on the development of vision systems for markerless motion capture for biomechanical applications currently lies on employing passive systems. Passive systems are advantageous as they only rely on capturing images and thus provide an ideal framework for capturing subject in their natural environment.

Over the past two decades, the field of registering human body motion using computer vision has grown substantially. Many examples of tracking and estimating human motion using models of different kinds have been proposed [15-35]. These are typically classified into model-based and model-free approaches. For example, model-based involve an a priori model which has relevant anatomic and kinematic information and is tracked or matched to 2D images [17, 19] or 3D shapes [20]. Moreover, several surveys concerned with computer-vision approaches have been published in recent years, each classifying existing methods into different categories [36-40]. For instance, Moeslund et al. [39] reviewed more than 130 human motion capture papers published between 1980 and 2000. The field was not categorized by approaches or techniques used, but rather by the stages necessary to solve the general problem of motion capture. Moeslund et al. categorized motion capture approaches into: initialization, tracking, pose estimation, and recognition. Wang et. al [40] provided a similar survey of human motion capture approaches in the field of computer vision ranging mainly from 1997 to 2001 and including a wider review of motion segmentation and object classification.

The majority of research on human motion capture in the field of computer vision and machine learning has concentrated on tracking, estimation and recognition of human motion for surveillance purposes. In addition, much of the work reported in the literature on the above has been developed for the case when a single camera is available. Single image stream based methods suffer from poor performance for accurate movement analysis due to the severe ill-posed nature of motion recovery. Furthermore, simplistic models of a human body with either fewer joints or reduced number of degrees of freedom are often utilized for enhancing computational performance. Existing methods for gait-based human identification in surveillance applications use mostly 2D appearance models and measurements such as height, extracted from the side view. However, biomechanical and, in particular, clinical applications typically require knowledge of detailed and accurate representation of 3D joint mechanics. Some of the most challenging issues in whole-body movement capture are due to the complexity and variability of the appearance of the human body, the nonlinear and non-rigid nature of human motion, a lack of sufficient image cues about 3D body pose, including self-occlusion as well as the presence of other occluding objects, and exploitation of multiple image streams. Human body self-occlusion is a major cause of ambiguities in body part tracking using a single camera. The self-occlusion is addressed when multiple cameras are used, since the appearance of a human body from multiple viewpoints is available.

While existing computer vision approaches offer great potential for markerless motion capture for biomechanical applications, these approaches have not been developed or tested for biomechanical applications. To date, qualitative tests and visual inspections are most frequently used for assessing approaches introduced in the field of computer vision and machine learning. Thus, evaluating existing approaches within a framework focused on addressing biomechanical applications is critical. Quantitative measurements of movement and continuous tracking of humans using multiple image streams is crucial for 3D gait studies. Thus, our emphasis currently focuses on the development of motion capture methods for use with multiple synchronized cameras. To critically analyze the effectiveness of markerless motion capture in the biomechanical/clinical environment, we quantitatively evaluated it using comparison with real data obtained from marker-based motion capture. In addition, realistic biomechanical models were used for identifying human body kinematics.

3. METHODOLOGY

Subjects performed walking trials at a self-selected normal speed while being recorded simultaneously using a markered and a markerless motion capture system. The marker-based system consisted of an eight-Qualisys camera optoelectronic system monitoring 3D marker positions for the hip, knees and ankles at 120 fps. The markerless motion capture system consisted of eight Basler CCD color cameras synchronously capturing images at 75 fps. Similar to stereophotogrammetry, the configuration of the technical equipment and calibration is critical. A most favorable camera arrangement for a 3 x 1.5 x 2m viewing volume was used to provide optimal visual construction [41]. This viewing volume is sufficiently large enough to capture an entire gait cycle. Internal and external camera parameters and a common global frame of reference are obtained through offline calibration. Images from all cameras were streamed in their uncompressed form to several computers during acquisition (Figure 1).

Figure 1: Selected views of 4 of the 8 Basler CCD color cameras.

In this study, three different approaches for estimating human movement from multiple image sequences for biomechanical applications were explored. The first two approaches tracked a 3D articulated model in 3D representations constructed from the image sequences, while the third approach tracked a 3D articulated model in multiple 2D image planes. The articulated model is typically derived from a morphological description of the human body's anatomy plus a set of information regarding the kinematic chain and joint centers. The morphological information of the human body can be a general approximation (cylinders, super quadrics, etc.) or an estimation of the actual subject's outer surface. Ideally, an articulated model is subject-specific and created from a direct measurement of the subject's outer surface. The kinematic chain underneath an anatomic model can be manually set or estimated through either functional [16, 18] or anthropometric methods [42]. The more complex the kinematic description of the body the more information can be obtained from the 3D representation matched by the model. In the current protocol, a detailed articulated body is created from a full body laser scan with markers affixed to the subject's joints (Figure 2). The articulated body consisted of 15 body segments and 14 joints connecting these segments.

a) b) c)

Figure 2: a) 3D surface from laser scan. b) Body segments. c) Articulated model.

Tracking a 3D articulated model in 3D representations

In the first two approaches movement was determined by first constructing 3D representations of the subject in the form of visual hulls and subsequently labeling individual body segments for extracting human body kinematics. The subject was separated from the background in the image sequence of all cameras using intensity and color thresholding [43] compared to background images (Figure 3). Segmentation is a challenging task as changes in illumination and shadow make quick and reliable processing difficult. Numerous approaches for segmentation have been used in indoor and outdoor settings. Most current segmentation methods use either temporal or spatial information of images including background subtraction, statistical methods, temporal differencing, and optical flow (see Wang [40] for review). Background subtraction is most frequently used for segmentation in situations with a relatively static background, and thus seems promising for a laboratory setting. However, this approach is very sensitive to changes in the background image such as illumination and may not be appropriate for all targeted applications.

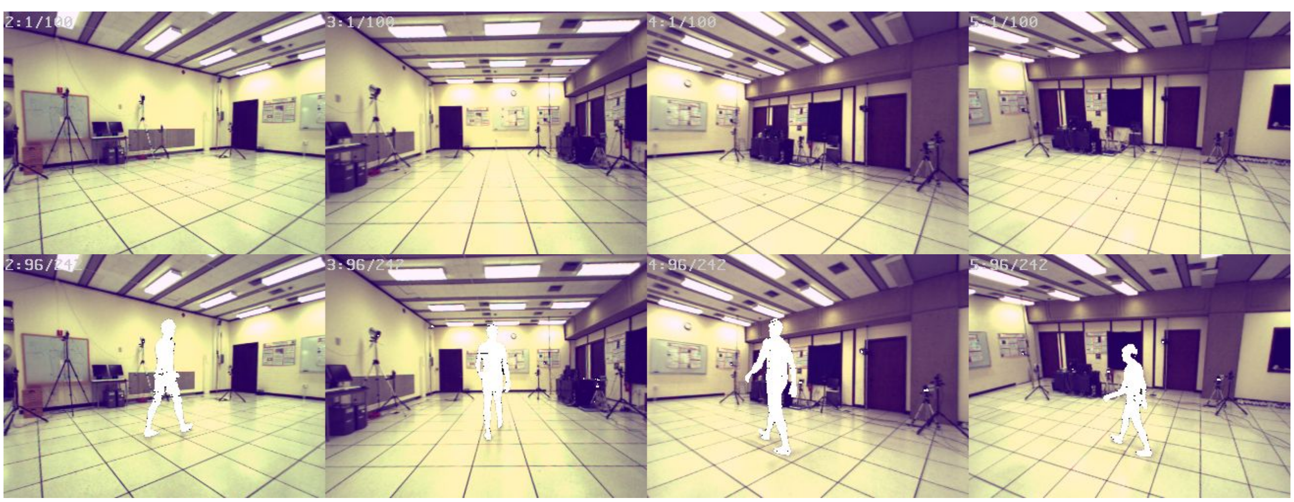

Figure 3: Selected background images (top) and separated subject data (bottom).

A standard way of capturing dynamic 3D sequences from images is visual hull construction which produces surfaces of lesser quality. The 3D representation was achieved through visual hull construction from multiple 2D camera views [44-46]. The number of cameras used for visual hull construction greatly affects the accuracy of visual hulls [41]. Visual hulls were created with voxel edges of λ = 10 mm (Figure 4), which is sufficiently small enough for these camera configurations [47]. The representation of the subject in the form of visual hulls offers an advantageous suitability for numerous estimation approaches. For example, the 3D surfaces can also be utilized for anthropometric measurements such as body segment volumes [47].

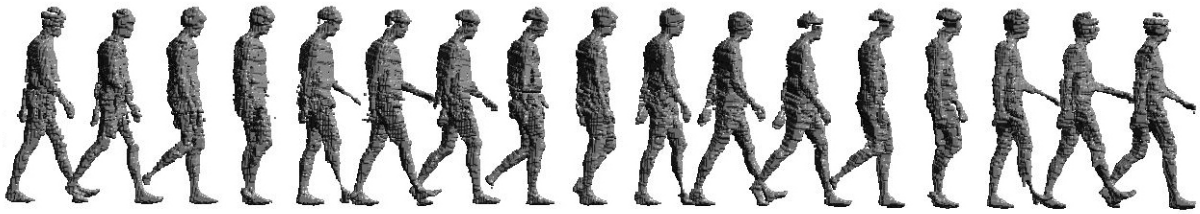

Figure 4: Sequence of visual hulls constructed using a passive vision system.

An articulated model (Figure 2) was roughly matched to a 3D representation in the motion sequence (Figure 4) and subsequently tracked automatically. Two different approaches were evaluated. The first approach is a gradient based articulated iterative closest point (ICP) algorithm with soft joint constraints [48]. This algorithm is a generalization of the standard ICP algorithm [49] to articulated models. The second approach is an adapted fast simulated annealing algorithm running on an exponential maps geometry formulation [42, 50]. This algorithm is a statistical computational method based on Boltzmann Sampling and the Metropolis Monte Carlo method. Joint angles for the sagittal and frontal plane for the knee calculated as angles between corresponding axes of neighboring segments were used as preliminary basis of comparison between the marker-based and markerless systems.

Tracking a 3D articulated model in multiple 2D image planes

The third approach extracted structure and motion through optical flow. The underlying algorithm as proposed in [33] combines multiple cues, such as pixel displacements, silhouettes and "motion residues" to track the pose. The objective was to estimate the pose at the current time instant *t*, given the pose at the previous time instant *t-1*, and the images at the previous and current time instants. The approach used motion information measured as pixel displacements and spatial information that included "motion residues" which described the boundary of the body segments as well as 2D image silhouettes in the tracking algorithm. Motion and spatial cues are complementary to each other. The 2D pixel displacement between frames was computed for each body segment and each image. The pixel displacement is a function of the change in pose, and can be used to predict the pose at the current time instant given the pose at the previous time instant. Using only the motion information in the tracking causes the tracking error to accumulate over time and leads to drift. Therefore, the pose was corrected at each time instant using the spatial cues. This method thus combines motion and spatial cues in a predictor-corrector framework.

The pose at the current time instant was estimated by first projecting the pose at the previous time instant onto each image (Figure 5). Pixels in each image were registered to the corresponding body segment, while the corresponding body segments provide 3D coordinate for some pixels in the image planes. The 3D coordinates for the remaining pixels were obtained through interpolation.

a) b) c) d)

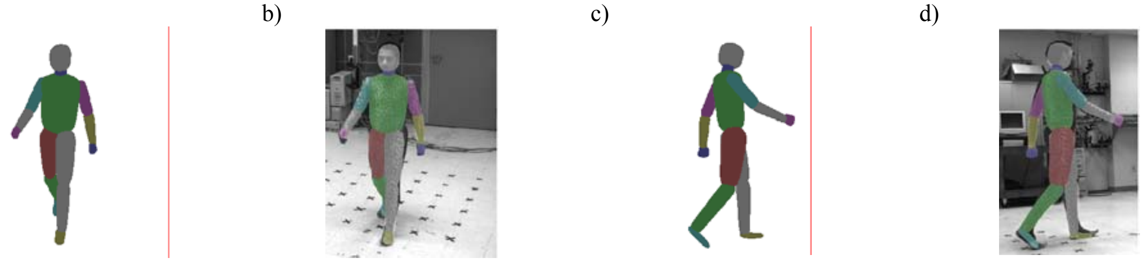

Figure 5: The 3D model of the object projected onto the image. Each pixel is registered to a body segment.

Next, pixel displacement between the previous and current time instants for all pixels in the mask for each body segment and each image was estimated. The pixel displacement was computed using optical flow at multiple scales and a rigid body model for the pixel displacement. The image at the previous time instant can be warped according to the estimated motion and the difference between the current image and the warped previous image is the motion residue (Figure 6). The pose at current time instant was predicted using pixel displacement of pixels belonging to all body segments and in all images. The pixel velocity is a linear function of the pose [32] and the change in body pose can be estimated iteratively using the measured pixel displacement.

(a) Image at *t-1* (b) Image at *t* (c) Mask (d) Pixel displacement (e) Motion residue

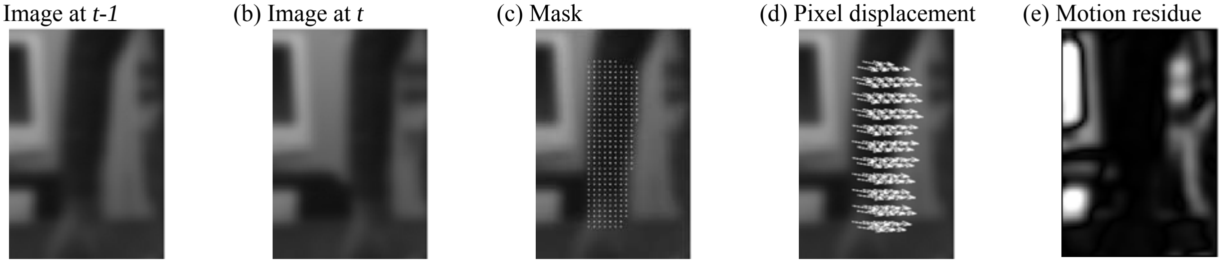

Figure 6: Computing pixel displacement and "motion residue".

Finally, silhouette and "motion residue" (Figure 7) are combined for each body segment into an "energy image" for each body segment and each image. An error function of the pose at time *t* is computed using these energy images. The pose is refined by minimizing the error function which is dependent on the pose. The predicted pose is corrected at time *t* using the "energy image" (Figure 7) obtained in previous step using optimization.

a) Silhouette of subject b) Silhouette c) Motion residue d) Energy image

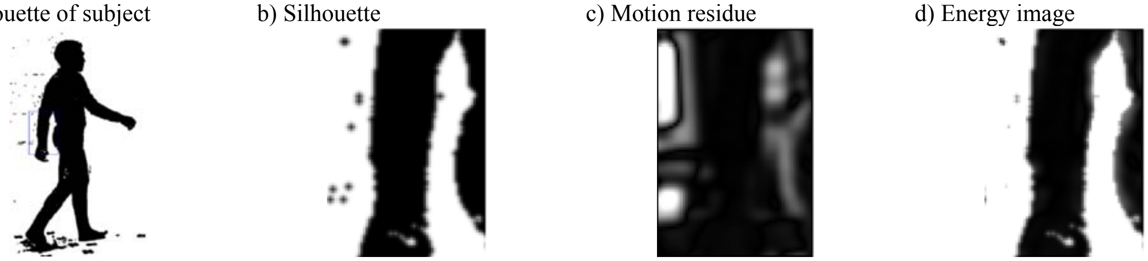

Figure 7: Combining spatial cues to obtain spatial energy image.

4. RESULTS

For the first two approaches, the 3D representation was achieved through visual hull construction from multiple 2D camera views. The number of cameras used for visual hull construction greatly affects the accuracy of visual hulls [41]. Surface comparison between visual hulls and the original human form revealed under-approximated and over-approximated regions. Under-approximated regions result from discretization errors in the image plane, which can be reduced with higher imager resolution [47]. However, greater error arises from over-approximated regions, which are characteristic to visual hull construction. The size of over-approximated regions and the maximum surface deviations decreases with increasing number of cameras. The accuracy of visual hulls also depends on the human subject's position and pose within an investigated/observed viewing volume [41]. Simultaneous changes in position and pose result in decreased accuracy of visual hull construction. Increasing the number of cameras leads to decreased variations across the viewing volume and a better approximation of the true volume value.

In general, body segments were tracked accurately with both approaches (Figure 8). The accuracy of sagittal and frontal plane knee joint angles calculated from experiments was similar to the accuracy estimated from theoretical calculations [51, 52] (Table 1) and match the accuracy of marker-based systems [53].

Table 1: Accuracy of sagittal and frontal plane kinematics for the knee of the right leg.

	theoretical		experimental	
	sagittal plane	frontal plane	sagittal plane	frontal plane
articulated ICP	2.1 ± 0.9°	0.4 ± 0.7°	2.3 ± 1.0°	1.6 ± 0.9°
simulated annealing	1.5 ± 3.9°	2.0 ± 2.3°	1.4 ± 4.3°	3.2 ± 6.0°

Figure 8: (left) Visual hulls. (right) Articulated body matched to visual hulls.

For the third approach, image sequences were tracked using both spatial and motion cues. Two sequences of a subject performing different activities were tracked. The length of the first sequence is 10 seconds (300 frames), during which there is considerable movement and bending of the arms and occlusions at various times in different cameras. The second sequence is that of the subject walking and is approximately 3 seconds long. The use of motion cues only resulted in poor pose estimation. The algorithm ultimately lost track due to error accumulation. This underlines the need for "correcting" the pose estimated using only motion cues. It was observed that the "correction" step of the algorithm prevented drift in the tracking. The body parts were successfully tracked in both cases. Figure 9 shows results in which the model assuming the estimated pose was superimposed over the images for two cameras.

a) b)

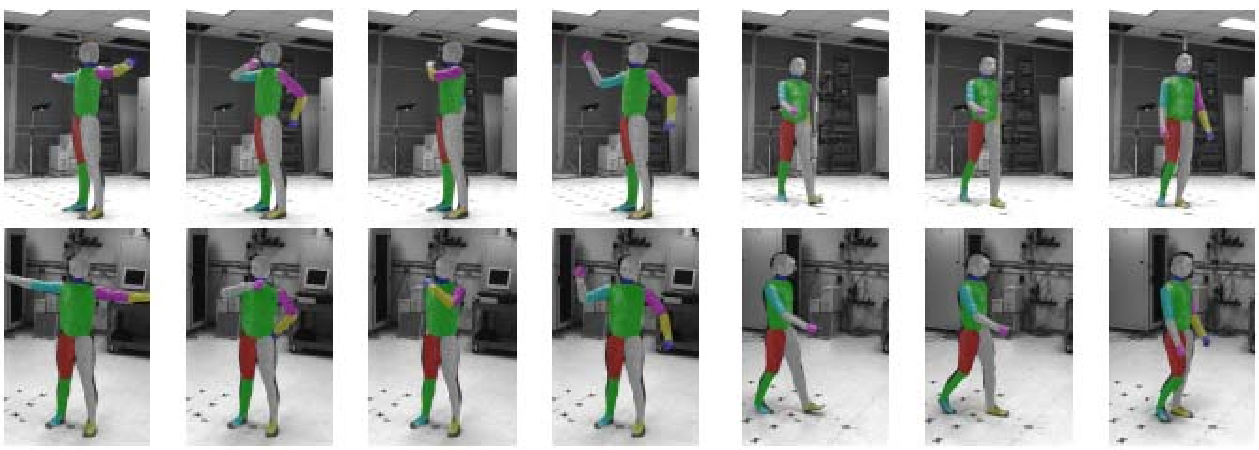

Figure 9: Tracking results for two sequences.

5. DISCUSSION

The development of markerless motion capture methods presented in this document was motivated by the need to address contemporary needs to understand normal and pathological human movement without the encumbrance of markers or fixtures placed on the subject, while achieving the quantitative accuracy of marker based systems. Markerless motion capture has been widely used for a range of applications in the surveillance, film and game industries. However, the biomechanical, medical, and sports applications of markerless capture have been limited by the accuracy of current methods for markerless motions capture. Previous experience has demonstrated that minor changes in patterns of locomotion can have a profound impact on the outcome of treatment or progression of musculoskeletal pathology. The ability to address emerging clinical questions on problems that influence normal patterns of locomotion requires new methods that would limit the risk of producing artifact due to markers or the constraints of the testing methods.

This study demonstrates the feasibility of accurately and precisely measuring 3D human body kinematics using a markerless motion capture system. The human body kinematics from the markerless and marker-based system produced comparable results. The errors affecting the accuracy of a markerless motion capture system can be classified into errors due to limitations of the technical equipment and errors due to the shape and/or size of the object or body under investigation. Another limitation of current methodology is its limited ability to resolve skeletal motions from markers placed on the skin. A better understanding of skeletal motion is required for the next productive step toward understanding musculoskeletal function. Current marker-based systems sample the motions of relatively few points located on the surface of the skin, or worse, use fixtures tightly strapped around the segment to estimate its motion. The approaches offer the possibility of reducing error by measuring the motion and shape of the entire limb segment.

Accuracy of markerless methods based on visual hulls is dependent on the number of cameras. In general, configurations with fewer than 8 cameras yielded two main drawbacks. First, different camera placements yielded different results for volume estimations and coefficient of variation. Second, volume estimations greatly deviated from original values and fluctuated enormously for different poses and positions across the viewing volume. Thus, configurations with less than 8 cameras will not capture human movement using visual hulls with sufficient accuracy for most biomechanical applications. However, configurations with 8 and more cameras provided good volume estimations and consistent results for different poses and positions across the viewing volume.

The markerless framework established in this work can serve as a basis for developing the broader application of markerless motion capture. Each of the modules can be independently evaluated and modified as newer methods become available. Markerless motion capture systems offer the promise of expanding the applicability of human movement capture, minimizing patient preparation time, and reducing experimental artifacts caused by inter-operator variability in marker placement, skin deformation, and the sensation of the applied markers.

ACKNOWLEDGEMENTS

This project was funded by NSF grant #03225715 and VA grant #ADR0001129.

REFERENCES

1. Murray, M.P., A.B. Drought, and R.C. Kory, *Walking patterns of normal men.* Bone and Joint Surgery, 1964. **46**(A): p. 335-360.
2. Murray, M.P., R.C. Kory, and S.B. Sepic, *Walking patterns of normal women.* Archives of Physical Medicine and Rehabilation, 1970. **51**: p. 637-650.
3. Johannson, G., *Visual perception of biological motion and a model for its analysis.* Perception and Psychophysics, 1973. **14**: p. 201-211.
4. Cutting, J. and L. Kozlowski, *Recognizing friends by their walk: gait perception without familiarity cues.* Bulletin of the Psychonomic Society, 1977. **9**: p. 353-356.
5. Wong, M.A., S. Simon, and R.A. Olshen, *Statistical analysis of gait patterns with cerebral palsy.* Statistics in medicine, 1983. **2**(345-354).
6. Winter, D.A., *Biomechanical motor patterns in normal walking.* Journal of Motor Behaviour, 1983. **15**: p. 302-330.
7. Holzreiter, S.H. and M.E. Kohle, *Assessment of gait patterns using neural networks.* 1995. **26**: p. 645-651.
8. Weber, W. and E. Weber, *Mechanik der menschlichen Gehwerkzeuge*. 1836, Göttingen: Dieterich.
9. Marey, E., *Animal Mechanism: A Treatise on Terrestrial and Aerial Locomotion*. 1874, London: Henry S. King & Co.
10. Muybridge, E., *Animal locomotion*. 1887, Philadelphia: J.B. Lippincott Company.
11. Andriacchi, T.P., *Dynamics of pathological motion: Applied to the anterior cruciate deficient knee.* Journal of Biomechanics, 1992. **23**(1): p. 99-105.
12. Prodromos, C.C., T.P. Andriacchi, and J.O. Galante, *A relationship between gait and clinical changes following high tibial osteotomy.* Journal of Bone and Joint Surgery, 1985. **67**(8): p. 1188-1194.
13. Andriacchi, T.P. and R.P. Mikosz, *Basic science of the knee and total knee arthroplasty.* Current Science, 1992. **3**: p. 13-17.

14. Sharma, L., et al., *Knee adduction moment, serum hyaluronan level, and disease severity in medial tibiofemoral osteoarthritis.* Arthritis and Rheumatism, 1998. **41**(7): p. 1233-1240.
15. Deutscher, J., A. Blake, and I. Reid. *Articulated body motion capture by annealed particle filtering*. in *Computer Vision and Pattern Recognition*. 2000. Hilton Head, SC.
16. Bottino, A. and A. Laurentini, *A silhouette based technique for the reconstruction of human movement.* Computer Vision and Image Understanding, 2001. **83**: p. 79-95.
17. Bregler, C., J. Malik, and K. Pullen, *Twist based Acquisition and Tracking of Animal and Human Kinematics.* International Journal of Computer Vision, 2004. **56**(3): p. 179-194.
18. Cheung, G., S. Baker, and T. Kanade, *Shape-From-Silhouette Across Time Part II: Applications to Human Modeling and Markerless Motion Tracking.* International Journal of Computer Vision, 2005. **63**(3): p. 225-245.
19. Isard, M. and A. Blake. *Visual tracking by stochastic propagation of conditional density*. in *4th European Conference on Computer Vision*. 1996. Cambridge, UK.
20. Gavrila, D. and L. Davis. *3-D model-based tracking of humans in action: a multi-view approach*. in *Conference on Computer Vision and Pattern Recognition*. 1996. San Francisco, CA.
21. Metaxes, D. and D. Terzopoulus, *Shape and nonrigid motion estimation through physics-based synthesis.* IEEE Trans. Patt. Anal. Mach. Intell., 1993. **15**: p. 580-591.
22. Kakadiaris, I.A. and D. Metaxes, *3D human body model acquisiton from multiple views.* Intl. Jl. Computer Vision, 1998. **30**: p. 191-218.
23. Kakadiaris, I.A. and D. Metaxes, *Model-based estimation of human motion.* IEEE Trans. Patt. Anal. Mach. Intell., 2000. **22**: p. 1453-1460.
24. Black, M.J., et al. *Learning parameterized models of image motion*. in *Computer Vision and Pattern Recognition*. 1997. San Juan, Puerto Rico.
25. Kanade, T., et al. *Advances in co-operative multi-sensor video surveillance*. in *DARPA Image Understanding Workshop*. 1998.
26. Cheng, J. and J.M.F. Moura, *Capture and representation of human walking in live video sequences.* IEEE Trans. on Multimedia, 1999. **1**: p. 144-156.
27. Stauffer, C. and W.E.L. Grimson. *Adaptive background mixture models for real-time tracking*. in *Computer Vision and Pattern Recognition*. 1999. Fort Collins, CO.
28. Krumm, J., et al. *Multi-camera multiperson tracking for EasyLiving*. in *3rd IEEE International Workshop on Visual Surveillance*. 2000. Dublin, Ireland.
29. Sidenbladh, H., M.J. Black, and D.J. Fleet. *Stochastic tracking of 3D human figures using 2D image motion*. in *European Conference on Computer Vision*. 2000. Dublin, Ireland.
30. Lee, J., et al., *Interactive Control of Avatars Animated with Human Motion Data.* ACM Transaction on Graphics, 2002. **21**(3): p. 491-500.
31. Moon, H., R. Chellappa, and A. Rosenfeld. *3D object tracking using shape-encoded particle propagation*. in *Intl. Conf. on Computer Vision*. 2001. Vancouver, BC.
32. Sundaresan, A., A. RoyChowdhury, and R. Chellappa. *Multiple view tracking of human motion modelled by kinematic chains*. in *International Conference on Image Processing*. 2004. Singapore.
33. Sundaresan, A. and R. Chellappa. *Multi-camera tracking of articulated human motion using motion and shape*. in *Asian Conference on ComputerVision*. 2006. Hyderabad.
34. Wren, C.R., et al., *Pfinder: Real-time tracking of the human body.* Trans. on Pattern Analysis and Machine Intelligence, 1997. **19**(7): p. 780-785.
35. Cheung, G., S. Baker, and T. Kanade. *Shape-from-silhouette of articulated objects and its use for human body kinematics estimation and motion capture*. in *IEEE Conference on Computer Vision and Pattern Recognition*. 2003. Madison, WI: IEEE.
36. Cedras, C. and M. Shah, *Motion-based recognition: a survey.* Image and Vision Computing, 1995. **13**(2): p. 129-155.
37. Aggarwal, J. and Q. Cai, *Human motion analysis: a review.* Computer Vision and Image Understanding, 1999. **73**(3): p. 295-304.
38. Gavrila, D., *The visual analysis of human movement: a survey.* Computer Vision and Image Understanding, 1999. **73**(3): p. 82-98.
39. Moeslund, G. and E. Granum, *A survey of computer vision-based human motion capture.* Computer Vision and Image Understanding, 2001. **81**(3): p. 231-268.

40. Wang, L., W. Hu, and T. Tan, *Recent Developments in Human Motion Analysis.* Pattern Recognition, 2003. **36**(3): p. 585-601.
41. Mündermann, L., et al., *Most favorable camera configuration for a shape-from-silhouette markerless motion capture system for biomechanical analysis.* SPIE-IS&T Electronic Imaging, 2005. **5665**: p. 278-287.
42. Corazza, S. and C. Cobelli. *An accurate model-based approach for markerless motion capture*. in *Medicon*. 2004. Naples, Italy.
43. Haritaoglu, I. and L. Davis, *W4: real-time surveillance of people and their activities.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. **22**(8): p. 809-830.
44. Martin, W. and J. Aggarwal, *Volumetric description of objects from multiple views.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1983. **5**(2): p. 150-158.
45. Laurentini, A., *The Visual Hull concept for silhouette base image understanding.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994. **16**: p. 150-162.
46. Cheung, K., S. Baker, and T. Kanade, *Shape-From-Silhouette Across Time Part I: Theory and Algorithm.* International Journal of Computer Vision, 2005. **62**(3): p. 221-247.
47. Mündermann, L., et al., *Conditions that influence the accuracy of anthropometric parameter estimation for human body segments using shape-from-silhouette.* SPIE-IS&T Electronic Imaging, 2005. **5665**: p. 268-277.
48. Anguelov, D., L. Mündermann, and S. Corazza. *An Iterative Closest Point Algorithm for Tracking Articulated Models in 3D Range Scans*. in *Summer Bioengineering Conference*. 2005. Vail, CO.
49. Besl, P. and N. McKay, *A method for registration of 3D shapes.* Transactions on Pattern Analysis and Machine Intelligence, 1992. **14**(2): p. 239-256.
50. Corazza, S., et al., *A markerless motion capture system to study musculoskeletal biomechanics: visual hull and simulated annealing approach.* Annals of Biomedical Engineering, Submitted, 2005.
51. Mündermann, L., et al. *Estimation of the accuracy and precision of 3D human body kinematics using markerless motion capture and articulated ICP*. in *Summer Bioengineering Conference*. 2005. Vail, CO.
52. Mündermann, L., et al. *Validation of a markerless motion capture system for the calculation of lower extremity kinematics*. in *International Society of Biomechanics & American Society of Biomechanics*. 2005. Cleveland, OH.
53. Chiari, L., et al., *Human movement analysis using stereophotogrammetry - part 2: Instrumental errors.* Gait and Posture, 2005. **21**(2): p. 197-211.