Computer Science Department of The University of Auckland
CITR at Tamaki Campus (http://www.citr.auckland.ac.nz)

CITR-TR-71 November 2000

A Comparison of Feature Measurements for Kinetic Studies on Human Bodies

Nikki Austin[2], Yen Chen[1], Reinhard Klette[1], Robert Marshall[2], Yuan-sheng Tsai[1], and Yongbao Zhang[1]

Abstract

The paper reports about a performance comparison within a joint project of computer vision, and sport and exercise sciences. The project is directed on the understanding of human motion based on shape features and kinetic studies. Three shape recovery techniques, a traditional technique as used in sport and exercise sciences (manual measurement based on an elliptical zone assumption) and two computer vision techniques (based on a small number of occluding contours, and a new combination of photometric stereo and shape from boundaries), are compared using a mannequin as test object. The computer vision techniques have been designed to go towards dynamic shape recovery (humans in motion). The paper reports about these three techniques and their measurement accuracies.

[1] Centre for Image Technology and Robotics, Computer Vision Unit, The University of Auckland, Auckland, New Zealand

[2] Department of Sport and Exercise Sciences, Tamaki Campus, The University of Auckland, Auckland, New Zealand

A Comparison of Feature Measurements for Kinetic Studies on Human Bodies

Nikki Austin[2], Yen Chen[1], Reinhard Klette[1], Robert Marshall[2], Yuan-sheng Tsai[1] and Yongbao Zhang[1]

[1]Center for Image Technology and Robotics
[2]Department of Sport and Exercise Sciences
Tamaki Campus, The University of Auckland, Auckland, New Zealand

Abstract. The paper reports about a performance comparison within a joint project of computer vision, and sport and exercise sciences. The project is directed on the understanding of human motion based on shape features and kinetic studies. Three shape recovery techniques, a traditional technique as used in sport and exercise sciences (manual measurement based on an elliptical zone assumption) and two computer vision techniques (based on a small number of occluding contours, and a new combination of photometric stereo and shape from boundaries), are compared using a mannequin as test object. The computer vision techniques have been designed to go towards dynamic shape recovery (humans in motion). The paper reports about these three techniques and their measurement accuracies.

1 Introduction

In sport science, biomedical engineering or ergonomics, accurate measurements of segmental anthropometry are often required for the estimation of resultant moments and forces acting on body segments. As it is not possible to directly measure the forces and moments acting on the human body, it is necessary to calculate them from measures of the external forces, the kinematic characteristics of the segment (linear and angular position, velocity and acceleration) and estimates of the segment's inertial parameters (center of mass location, segment length and segment moment of inertia).

We specify features of interest [11]. We assume an XYZ-cartesian world coordinate system with the Z-axis being the axis of gravitation, i.e. the weight acts in Z-direction. Consider a human body segment s and its local xyz-coordinate system with its origin at the center of mass of s, and with the z-axis aligned with the main ("long") axis of the segment. We assume and use uniform density estimates for human body segments [2, 3]. The forces acting on the center of mass are in world coordinates

$$\mathbf{F}_{j_1,s} \; + \; \mathbf{F}_{j_2,s} \; + (0,0,w_s) \; = \; m_s \cdot \mathbf{a}_s \; ,$$

where j_1, j_2 are the joints of the given segment s (note: there is only one joint for some segments), $\mathbf{F}_{j,s} = (F_X, F_Y, F_Z)$ are the forces acting on the segment

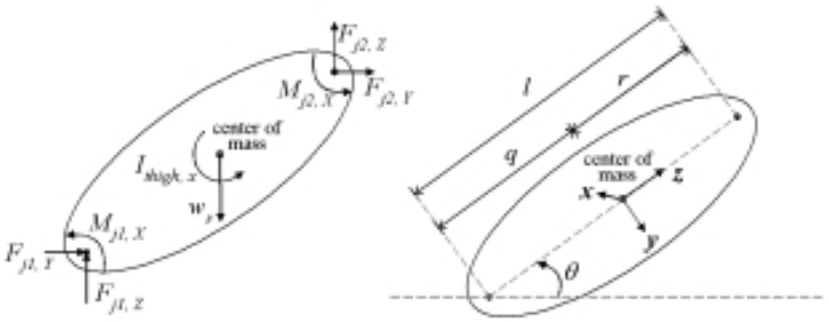


Fig. 1. Left: forces acting on a cross section of a thigh. Right: general scheme of the segment.

s at joint j, m_s is the mass of the segment s, and $\mathbf{a}_s = (a_X, a_Y, a_Z)$ are the accelerations of the segmental center of mass, all expressed with respect to the 3D world coordinate system. Figure 1 shows the forces acting on a two-dimensional cross section of the thigh, with $j_1 =$ knee, $j_2 =$ hip.

Kinematic studies require estimates of the *center of mass* (which requires shape data), of *accelerations* $\mathbf{a}_s$ (which requires video sequence analysis) and of m_s, i.e. of the *volume* of s using uniform segmental density estimates for deriving m_s. The location of the center of mass (in 2D specified by a value of r, see Fig. 1) is estimated based on shape analysis results in general, and an often used simplification is to assume that it is on the line of slope θ from joint to joint. The motion of the segment is assumed to be in the YZ-plane.

Furthermore, the moment of force in X-direction acting on the center of mass is [12]

$$M_{j_1,X} + M_{j_2,X} + q \cdot \sin\theta \cdot F_{j_1,Y} - q \cdot \cos\theta \cdot F_{j_1,Z} - r \cdot \sin\theta \cdot F_{j_2,Y} + r \cdot \cos\theta \cdot F_{j_2,Z} = I_x \cdot \alpha_x \ ,$$

where $q = l - r$, $\mathbf{M}_j = (M_{j,X}, M_{j,Y}, M_{j,Z})$ are the moments of force acting at joint j (we are not using $M_{j,Y}, M_{j,Z}$), I_x, I_y, I_z are the values in the segmental moment of inertia tensor (I_y, I_z are not of interest), and $\alpha_s = (\alpha_x, \alpha_y, \alpha_z)$ is the angular acceleration of the segment. The *base length* of the segment is given by the distance l between both joints, and r represents the distance from the center of mass to one of these joints, θ is the angle made by the line connecting the joints, and the horizontal axis. Kinematic studies require estimates of the *segmental moment of inertia* I_x and of the *segmental angular acceleration* α_x.

Accurate estimations of the listed parameters are essential to accurate calculations of forces and moments of forces. In this paper we only discuss estimates of the static parameters (center of mass, of the volume, and of the moments of inertia). The dynamic parameters (acceleration, angular acceleration) will

be discussed in another publication. Traditional techniques for estimating these static parameters as used in sport and exercise sciences include proportional estimates from cadaver measures, regression equations, mathematical modelling and scanning/imaging (MRI) techniques.

We also make use of a standard simplification of the model assuming that segments are only considered in a 2D YZ-coordinate system as shown in Fig. 1, and moments of force are considered with respect to rotation about the x-axis only.

Automatic computer-vision based whole body reconstruction systems that are currently available include a whole body color 3D scanner WB4, developed by Cyberware [4], a phase based body measurement system, developed by TC^2 [13], and a PC cluster system, under development at Kyoto university [14].

Cyberware's system uses structured lighting with four scanning instruments mounted on two vertical towers. The scanning instruments start scanning from the person's head and continue down to capture the shape and color of the human body. Recently such a scan requires about 17 seconds. The system is priced at US$410,000 each, excluding the graphical workstations required to run the system. TC^2's system uses a phase based approach to recover surface information. Six stationary sensors capture different light patterns that are projected onto the body. Recently a scan requires 8 seconds and the system is priced at US$100,000. The cited PC cluster system uses the shape from contours approach for shape reconstruction and is intended to achieve the reconstruction of a human body in real time. The cluster system includes 9 cameras and 10 PCs with dual Pentium-III 600 MHz CPUs and 256MB memories, connected by a high speed network.

The paper is structured as follows. Section 2 describes two methods for measuring the static parameters. Section 3 compares the results obtained by both methods.

2 Methods

For our performance test we choose the elliptical zone method [7] which is still in use for manual shape recovery, shape from contours (with a limitation to just a small number of contours, say 9), and a new computer-vision based method which combines photometric stereo based 2.5D shape data and an even smaller number of captured contours (say 3 or 4) to recover a full 3D shape. The motivation is that these methods allow dynamic 3D shape estimation due to possible frequencies. However, we have to answer the question up to what accuracy. We compare these techniques with the traditional elliptical zones method. True volume data may be obtained by water displacement measurement. A mannequin is used as a test object which is positioned on a turntable for both computer vision approaches (allowing that only one camera has to be used).

2.1 Elliptical zones

The elliptical zone technique [7] is a traditional technique in sport and exercise sciences. It considers each segment to be composed of a sequence of right *elliptical cylinders* e that follow the shape fluctuations of the segment. High resolution digital images from the front and side of a body are taken. For the segment under consideration assume that the side view corresponds to a projection into the YZ-plane.

An interactive program allows to divide a segment s into elliptical cylinders e. The two axes of such an elliptical cylinder, x_e, y_e, are measured in x- and y-direction of the segment, and the height h_e (which is typically about 2 cm) in z-direction. The volume of an elliptical cylinder e is estimated by

$$v_e \;=\; \pi \cdot x_e \cdot y_e \cdot h_e \;,$$

and the volume of a segment is simply the sum of the volumes of all of its elliptical cylinders. The mass m_e of this elliptical cylinder is its volume times its density (which is assumed to be uniform).

The center of mass of a segment s with n elliptical cylinders is estimated as follows: we assume that all centers of mass of all the elliptical cylinders are on the z-axis of the segment. The 50%-percentile specifies the center of mass of the segment, calculated by adding m_e-values along the segment's z-axis using the value l.

The moments of inertia of the base b (i.e. that's a planar region in 3D space) of such an elliptical cylinder e about its centroidal axes are denoted by $I_{x,b}$, $I_{y,b}$, $I_{z,b}$, and they are estimated by

$$I_{x,b} = \frac{\pi}{4} \cdot x_e \cdot y_e^3 \;, \quad I_{y,b} = \frac{\pi}{4} \cdot x_e^3 \cdot y_e \;, \quad \text{and} \quad I_{z,b} = I_{x,b} + I_{y,b} \;.$$

The moments of inertia of the 3D cylinder e about its centroidal axes are

$$I_{x,e} = I_{x,b} \cdot \rho_e \cdot h_e \;, \quad I_{y,e} = I_{y,b} \cdot \rho_e \cdot h_e \;, \quad \text{and} \quad I_{z,e} = I_{z,b} \cdot \rho_e \cdot h_e \;,$$

where ρ_e is the cylinder's (uniform) density. For the entire segment, the moments of inertia about its centroidal axes are defined with respect to its local xyz-coordinate system, and they are found by applying the parallel axes theorem and summing for $e = 1, ..., n$:

$$I_{x,s} = \sum_{e=1}^{n} (I_{x,e} + m_e \cdot d_e^2) \;,$$

where d_e is the distance between the center of mass (centroid) of the eth elliptical cylinder and the segmental centroid. Due to our assumption that the cylinder centroids are located on the z-axis it follows that the summed cylinder centroid inertia tensor is on the principal segmental axes.

2.2 Shape from photometric stereo and contours

Shape from contours is a method that obtains a 3D model from the occluding contours of an object [10]. The shape from contours approach is a robust method that gives reliable 3D shape estimation. Nevertheless, due to the nature of the approach, surface cavities that are obstructed from the viewing directions by other regions on the surface are unable to be recovered, and normally many contours (say 80 ... 150) are used to complete a 3D shape scan. We will limit our approach to a small number of contours (viewing directions). We use Tsai's calibration method [15, 16] for obtaining intrinsic and extrinsic camera parameters.

Furthermore we use a new combination of photometric stereo and shape from contours. The photometric stereo approach for surface recovery calculates local surface orientations according to surface irradiance values [6, 8, 17]. Local surface orientations are globally integrated to recover the surface depth values. The photometric stereo approach allows to recover 2.5D surfaces in real-time. However, the recovered surface depth values are relatively scaled.

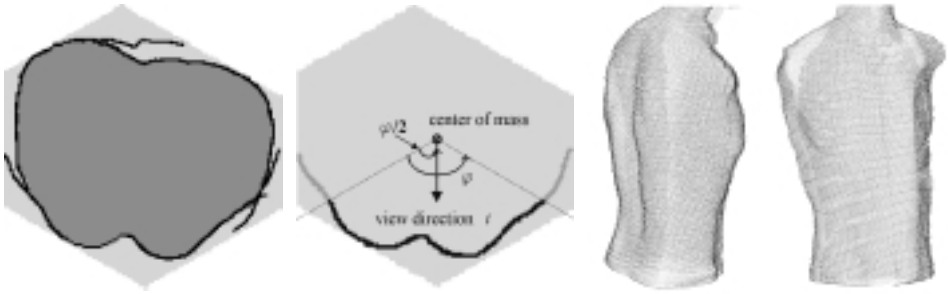
Three light sources successively illuminate the object from directions s_1, s_2 and s_3, and images E_{i1}, E_{i2} and E_{i3} are respectively acquired, where i is the index of current position. After three images have been acquired, the turntable is rotated by ϕ degrees to rotate the object into the next viewing direction, with the index of $i + 1$. The process of image acquisition is repeated for all of the required viewing directions, typically 3 or 4 only. Photometric stereo method is used to recover 2.5D surfaces from the three input images in any viewing direction.

Each layer e in the 3D image data has a height of 1 pixel, which is typically about 2.5 mm on the surface of the mannequin in our set-up. In any layer, all contours from all viewing directions define a convex polygon, as shown by the light gray region in Fig. 2(a). The surface pixels recovered by photometric stereo for the eth layer are fitted into the polygon according to the assumption that all surface pixels must lie within the polygon defined by the contours. The center of mass of the polygon is calculated and each viewing direction is specified as a vector starting from the center of mass. A surface pixel obtained at viewing direction i is accepted if it lies within $\phi/2$ from either side of the viewing direction. Figure 2(b) shows the accepted surface pixels in black. The rejected pixels, which lie outside of the $\phi/2$ threshold, are shown in gray. The accepted surface pixels form a 3D model of the object, as shown by Fig. 2(c).

A region R_e is defined by the accepted surface pixels in one layer as indicated by the dark gray region in Fig. 2(a). The number of pixels that lie within R_e specifies the area of the region, a_e.

The volume of the object at the eth cross section v_e, is estimated by counting the number of pixels that belongs to a_e and multiplied by the pixel to metric ratio k (approximately 2.5 mm in our set-up)

$$v_e = k^3 \cdot a_e \ .$$



(a) (b) (c)

Fig. 2. (a), (b) A layer of the recovered mannequin torso and (c) reconstructed 2.5D surfaces of mannequin torso.

The mass of the cross section is its volume multiplied by its density ρ_e, as for the elliptical zones approach. The center of mass (c_x, c_y) for the eth layer is

$$c_x = \frac{1}{a_e} \sum_{(x,y)\in R_e} x \ , \quad c_y = \frac{1}{a_e} \sum_{(x,y)\in R_e} y \ .$$

The segment's center of mass along its z-axis is specified as the 50%-percentile of the cumulative sum of the cross section masses.

The moment of inertia for the base of a cross section about its center of mass is

$$I_{x,b} = k^2 \cdot a_e \cdot \sum_{(x,y)\in R_e} (x - c_x)^2, \ I_{y,b} = k^2 \cdot a_e \cdot \sum_{(x,y)\in R_e} (y - c_y)^2, \text{ and } \ I_{z,b} = I_{x,b} + I_{y,b} \ .$$

The moment of inertia for the entire segment about its centroidal axes are as given for the elliptical zones approach.

3 Results

In the situation where mass cannot be measured directly, accurate volume measurement is crucial to the estimation of mass, since mass is obtained by multiplying volume by density. The center of mass and moment of inertia can then be calculated once the mass for any layer of the segment can be estimated.

In this section, the segmental features calculated from different shape recovery techniques are compared with measurements that are obtained by physically measuring the object. Different approaches are discussed with respect to the applicability, performance and possible improvement.

3.1 Comparisons

The measurements are compared with measurements obtained by water displacement, balancing, and pendulum approaches. It is assumed that the density

is uniform over the entire object. The values for center of mass are given as distances from the apex of the head. The values of the moment of inertia are calculated with respect to a horizontal axis that passes through the center of mass.

Method	Result (L)	Error (%)
Elliptical zone	30.01	8.56%
Shape from contours (9 views)	32.90	0.24%
Shape from PS and contours (3 view)	27.30	16.82%

Table 1: Resultant volume measurements and relative percentage error, where the volume obtained by water displacement is 32.82 L.

Method	Result (m)	Error (%)
Elliptical zone	0.35	16.67%
Shape from contours (9 views)	0.45	7.14%
Shape from PS and contours (3 views)	0.45	7.14%

Table 2: Resultant center of mass measurements and relative percentage error, where the center of mass obtained by balance is 0.42 m from the apex.

Method	Result (kgm^2)
Elliptical zone	1.33
Shape from contours (9 views)	1.31
Shape from PS and contours (3 views)	0.87

Table 3: Calculated moment of inertia, where the moment of inertia obtained by pendulum is 0.49 kgm^2. This points out that the density of the mannequin is actually not uniform, which leads to calculated results that do not correspond well with the measurement.

3.2 Discussion

From the results it can be seen that shape from contours with 9 viewing directions has provided the closest estimated values to the reference volume, as well as for the reference center of mass. The reason is that shape from contours is robust and reliable for obtaining the 3D shape of the object. Nevertheless, the accuracy of this method is limited, since some cavity regions are irrecoverable from shape from contours alone.

The shape from photometric stereo and contours method has obtained an estimated value for the center of mass that is closest to the reference value. A cause of error in the experiment may be that the depth values obtained by photometric stereo have been incorrectly scaled, leading to larger cavities in the recovered 3D model and thus reducing the total volume of the object.

The traditional elliptical zones model approach has not obtained values that are closest to the reference values in any of the feature measurements. One

possible factor which limits the accuracy obtained by the elliptical zones model approach is the assumption that segment cross-sections are ellipsoidal.

Overall, the accuracy of the shape from photometric stereo and contours method is not limited by cavities on the the segments, nor assumptions of the segment shapes. The accuracy is dependent on the reflectance properties of the surface, the number of viewing directions and the resolution of the cameras. The 3D data can further be refined with control points to improve the accuracy of the proposed method.

In the experiment, the density is assumed to be uniform over the entire object for simplicity. In reality, the density is not uniform, and the assumption influences the accuracy for the calculated center of mass, which in turn, effects the calculated moment of inertia. Particularly in the mannequin, where the head is much denser than the torso. Thus the value obtained by physically measuring the moment of inertia does not correspond well to the calculated results under the uniform density assumption.

The ratio for converting from pixel to metric is assumed to be constant throughout the image data and in all x, y and z directions. Nevertheless, due to distortions in image acquisition, the ratio may be varied in different directions.

Water displacement, balancing and pendulum methods have been used to generate the ground truth for volume, center of mass and the moment of inertia. However, in practice, it is not feasible to immerse a person (or part of) in water to obtain the volume of body parts. It is also not possible to locate the center of mass by balancing the object, or obtain the moment of inertia by swinging the object. Furthermore, neither of the physical approaches, nor the elliptical zones method, is applicable for dynamic shape recovery. The shape from photometric stereo and contours method has very low time requirement for image acquisition and shape recovery. It is possible to achieve surface recovery with photometric stereo in real time. The reduction in time requirement reduces errors caused by the object's movement.

References

1. N. Austin: Estimation of segmental inertial parameter variation in children: Implication for gait analysis. Master thesis, Sport and Exercise Sciences, The University of Auckland (2000).
2. R.F. Chandler, C.E. Clauster, J.T. McConville, H.M. Reynolds, J.W. Young: Investigation of inertial properties of the human body. AMRL Technical Report, TR-74-137 (1975).
3. C.E. Clauster, J.T. McConville, J.W. Young: Weight, volume and center of mass of segments of the human body. AMRL Technical Report, TR-69-70 (1969).
4. Cyberware, Whole Body Color 3D Scanner: http://www.cyberware.com/products/index.html (1995).
5. W.D. Dempster: Space requirements of the seated operator. WADC Technical Report 55-159 (1955).
6. R. Frankot, R. Chellappa: A method for enforcing integrability in shape from shading algorithms. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **PAMI-10** (1988) 439–451.

7. R.K. Jensen: Estimation of the biomechanical properties of three body types using a photogrammetric method. *J. of Biomechanics* **11** (1978) 349-358.
8. R. Klette, K.Schlüns, *Height data from gradient fields*, Proc. SPIE, 2908, pp.204-215, 1996.
9. R. Klette, K. Schlüns, A. Koschan: *Computer Vision: Three-dimensional Data from Images*. Springer, Singapore (1998).
10. J. J. Koenderink: *Solid Shape*. Cambridge, MA, MIT Press (1990).
11. R.N. Marshall, R.K. Jensen, G.A. Wood: A general Newtonian simulation of an n-segment open chain model. *J. of Biomechanics* **18** (1985) 359–368.
12. B.M. Nigg, W. Herzog: *Biomechanics of the Musculo-Skeletal System*. Wiley, New York (1994).
13. Tc2, Body Measurement System: http://www.tc2.com/RD/RDBody.htm (1998).
14. S. Tokai, T. Wada, T. Matsuyama: Real time 3D shape reconstruction using PC cluster system. Proc. 3rd Internat. Workshop *Cooperative Distributed Vision*, Kyoto (1999) 171–187.
15. R.Y. Tsai: An efficient and accurate camera calibration technique for 3D machine vision. Proc. Internat. Conf. *Computer Vision and Pattern Recognition* (1986) 364–374.
16. Tsai Camera Calibration Software: http://www.ius.cs.cmu.edu/afs/cs.cmu.edu/user/rgw/www/TsaiCode.html (1995).
17. R.J. Woodham: Photometric method for determining surface orientation from multiple images. *Optical Engineering* **19** (1980) 139–144.