Statistical Estimation of Human Anthropometry from a Single Uncalibrated Image

Article · January 2008

CITATIONS
15

READS
127

2 authors, including:

Chiraz Benabdelkader
Ecole Nationale des Sciences de l'Informatique
27 PUBLICATIONS **1,365** CITATIONS

SEE PROFILE

Statistical Estimation of Human Anthropometry from a Single Uncalibrated Image

Chiraz BenAbdelkader[1] and Yaser Yacoob[2]

[1] New York Institute of Technology, Abu Dhabi Campus, Abu Dhabi, UAE
cbenabde@nyit.edu
[2] Institute of Advanced Computer Studies, University of Maryland, College Park, MD USA yaser@umiacs.umd.edu

1 Introduction

Anthropometric data refer to a collection of measurable physical dimensions and proportions of the human body. Anthropometrics have been used in a variety of fields; primarily in forensics and physical anthropology for quite a long time now, and more recently within the domains of apparel sizing and ergonomic workspace design, among others. With the advent and growth of computer-based image analysis, there is an increased interest in developing automated and semi-automated methods for estimating human anthropometry from images. Unlike traditional anthropometric measurement methods that perform the measurements directly on the person (sometimes using specialized instruments), image-based anthropometric methods are contactless; they obtain the measurements by processing images of the person. With high-resolution images from multiple calibrated views, these techniques are fairly simple and accurate [1]. But what if all we have is a single uncalibrated image; could any useful anthropometric information at all be recovered in this case? By *uncalibrated*, we mean that the conditions of image acquisition—the full camera calibration parameters—are unknown.

In the present work, we view an anthropometric as the (Euclidean) length of the 3D line segment between two well-defined and visible points on the human body, and we address the general problem of estimating anthropometrics from a single uncalibrated 2D image of a person. From a projective geometry point of view, this amounts to the basic 3D reconstruction problem of recovering the length of a 3D line segment based on its projection in a 2D image. In the calibrated case, this problem can be solved with additional knowledge of the segment's direction and a plane or line passing through one of its endpoints. In the uncalibrated case, a general approach for this problem consists of first computing the length ratio q of the line segment with

respect to another line segment in the scene of known length L, called the *reference length*, and then simply obtaining the solution as the product qL. This approach is purely geometric and comprises a whole family of metrology techniques [2, 3, 4, 5, 6]. However, it poses two main difficulties in practice: (i) a reference length may not always be available, and (ii) under the general perspective camera model, the length ratio q can only be computed under certain circumstances, such as when the two line segments are coplanar and the vanishing line of their plane is known, or when the two line segments are collinear and the vanishing point of their direction is known (See Section 3.4).

In this paper, we propose a novel metrology technique that resolves the first issue by obviating the need for the reference length altogether. Geometrically speaking, this is impossible because the reference length encodes absolute scale information that is essential for 3D reconstruction. Nonetheless, we will show that by injecting appropriate prior knowledge into the estimation process, an *approximate* solution for the problem can be obtained. Specifically, our technique simultaneously estimates the lengths of multiple (more than one) 3D line segments, based on their projections in a single image. First we compute their pairwise length ratios as in existing single-view metrology techniques, then we formulate constraints on the solution based on probabilistic/statistical knowledge about the unknown segment lengths and their ratios, assuming they vary statistically. This amounts to a linearly constrained quadratic function minimization problem, which we solve using standard quadratic programming techniques. We then apply this metrology technique for the estimation of **10 anthropometrics**, including head length and body height. The method has been tested both on synthetic data and real images—an inhouse dataset of 99 high-resolution images comprising 26 different adults from various ethnicities.

The main intuition for why this technique "works" lies in that human body dimensions and human body proportions vary within a relatively limited range, and so the projection of a person in an image encodes some information about the scale of things (i.e. absolute scale) in the imaged scene. Naturally, the technique is not as accurate as purely geometric techniques [3, 4], but this is the price to pay when working with insufficient information.

The rest of the Chapter is organized as follows. In Section 2, we review related work. In Section 3, we describe our novel metrology technique, which lays the foundation for estimating human anthropometrics in Section 4. In Section 5, we present the accuracy performance of our anthropometric estimation method. Finally, in Section 6, we conclude with a summary of the method, its achievements, and its limitations.

2 Related Work

Image-based anthropometric estimation, when done with multiple calibrated cameras is straightforward and requires hardly any attention from the research community, except when it comes to fully automating the entire process [1].

Anthropometric estimation in general has scarcely been addressed in the computer vision literature. However, certain anthropometrics have received some limited attention, namely facial anthropometrics (typically in the context of feature-based face recognition) and body height. Previous work on the estimation of body height from images varies in terms of whether the methods are automated or semi-automated, whether they use a single image or a sequence of images (video), and how much prior knowledge they assume about camera calibration and 3D structure of the scene [3, 7, 8, 9].

Another related body of work is that on the use of the human body as a camera self-calibration object. The work of [10, 11, 12] addressed the problem of camera self-calibration based on tracking a walking human in a video. It basically exploits the fact that a walking person is orthogonal to the ground plane and maintains the same height, with small variations during the walk cycle. Other work explored the use of human anthropometric proportions to constrain image understanding tasks [13, 14, 15].

To our knowledge, anthropometric estimation from single uncalibrated images has not been addressed in a comprehensive manner, as we do in this work. Our line of research is perhaps most akin to that of [16, 17]. In this work, Barron and Kakadiaris estimate pose and anthropometric ratios (and not anthropometrics per se) of a person in arbitrary pose from a single uncalibrated image. The proposed techniques are based on building a family of stick models and the of use of human anthropometry to constrain the estimation process. However, in our view, their main contribution really lies in body pose estimation and not in anthropometric estimation.

3 A Novel Image-Based Metrology Method

3.1 Overview

We describe a novel metrology technique that simultaneously estimates *multiple*, say n, 3D line segments (such as anthropometrics) from a single image under minimal calibration information, with $n > 1$. One common approach for this problem is as follows: first compute the ratio of two line segments based on their projections in the image, then compute the length of one segment as a function of the obtained ratio and the length of the other segment [4, 5, 6]. We extend this purely geometric approach by obviating the need for absolute length information and solving the problem in a statistical framework. For this, we assume the segment lengths to be estimated are random variables with known statistical properties. The missing absolute scale information (otherwise derived from the reference length) is now derived—indirectly speaking—from prior probabilistic/statistical knowledge. Figure 1 illustrates the relationship, in a nutshell, between our proposed technique and the existing purely geometric approach.

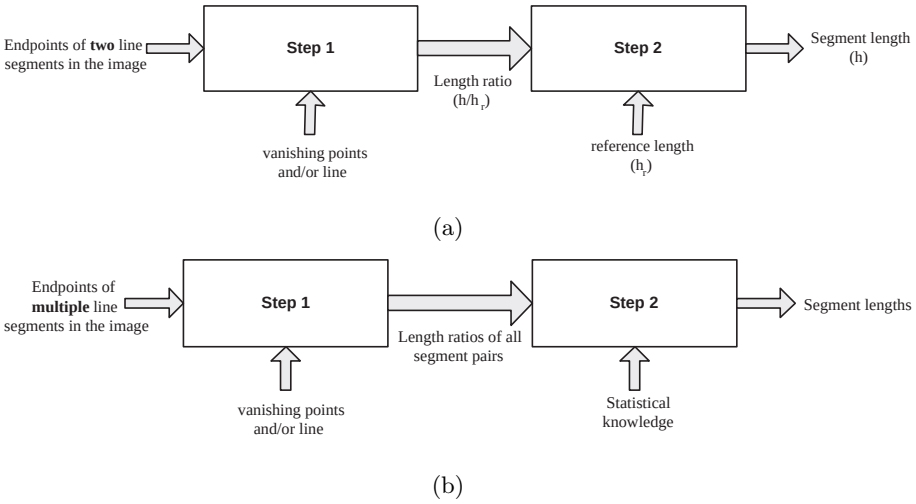


Fig. 1. Image-based metrology methods for computing the length of a 3D line segment under minimal calibration information; (a) Purely geometric existing techniques; (b) Our proposed technique, which extends the methods in (a) by obviating the need for absolute length information (the reference length) and incorporating prior knowledge of the statistical properties of the segments to be estimated.

3.2 An Ill-posed Inverse Problem Formulation

Let $\mathbf{x}$ be a vector containing the lengths of n 3D line segments, and let $\mathbf{y}$ be a vector containing m length ratios x_i/x_j where $i \neq j$. We wish to estimate $\mathbf{x}$ based on the observed value of $\mathbf{y}$. We will later go back to the issue of how the value of $\mathbf{y}$ is itself computed from a single image (Section 3.4).

From a purely algebraic point of view, this is an **ill-posed inverse problem**, as there are infinitely many vectors $\mathbf{x}$ that could generate the same value of $\mathbf{y}$; if a solution $\mathbf{x}_0$ exists, then *any* scalar multiple of $\mathbf{x}_0$ is also a solution. A formal derivation of the infinite solution space[3] for this problem follows.

Let $\mathcal{I}$ be the set of all pairs (i, j) such that x_i/x_j is in $\mathbf{y}$. Let r_{ij} be the observed value of the length ratio x_i/x_j, for any $(i, j) \in \mathcal{I}$. Each r_{ij} value provides a linear equality constraint in the unknown variables x_i and x_j:

$$x_i - r_{ij}x_j = 0 \tag{1}$$

Hence with m distinct length ratios we obtain m corresponding linear equality constraints, which we denote collectively by the linear homogeneous system:

$$\mathbf{Cx} = \mathbf{0} \tag{2}$$

[3] The term *solution space* of a problem in this paper is used to mean the set of candidate (admissible) solutions of the problem.

where $\mathbf{C}$ is hence a mxn matrix. Since r_{ij} and r_{ji} provide equivalent equality constraints (barring numerical errors), there can only be at most $n(n-1)/2$ distinct rows in $\mathbf{C}$. That is, $m \leq n(n-1)/2$.

We know from basic linear algebra that an exact nontrivial solution for the homogeneous linear system in (2) exists if and only if $\mathbf{C}$ is rank-deficient, i.e. $\text{rank}(\mathbf{C}) < n$. Otherwise, when $\text{rank}(\mathbf{C}) = n$, a non-exact nontrivial solution can be obtained by solving the following minimization problem via least squares estimation (LSE) [4]:

$$\min_{\mathbf{x}} \|\mathbf{C}\mathbf{x}\|^2 \qquad \text{subject to } \|\mathbf{x}\|^2 = 1 \tag{3}$$

The condition $\|\mathbf{x}\|^2 = 1$ serves both to avoid the trivial solution and to guarantee uniqueness. This LSE problem has a unique solution, which happens to be the right singular vector, denoted $\mathbf{v}_n$, of $\mathbf{C}$ that corresponds to its smallest singular value, denoted σ_n [18]. Thus, a set of approximate (inexact) solutions for (2) are:

$$\mathbf{x} = k\mathbf{v_n} \tag{4}$$

where k is an arbitrary scalar. This is an unbounded one-dimensional solution space, hence containing infinite candidate solutions. It is useful to notice that the sum of square errors (SSE) of each of these solutions is: $\|\mathbf{C}\mathbf{x}\|^2 = k^2\|\mathbf{C}\mathbf{v_n}\|^2 = k^2\sigma_n^2$, which is an increasing function of k. In other words, our infinite candidate solutions are not all equal in terms of their data fit, and the trivial solution solution has the best fit. Clearly, some other criterion is needed, in addition to data fit, for selecting a single "best" candidate solution from this infinite solution space.

3.3 Approximate Solution via Regularization

So far, we have derived a solution space of our inverse problem that is infinite and unbounded (Equation (4)), and we need a sensible way for selecting a single "best" solution from this infinitum of candidate solutions. A well-known strategy for dealing with ill-posed inverse problems is *regularization*—the process of introducing appropriate prior knowledge and assumptions about the unknown $\mathbf{x}$, in order to hopefully (but not necessarily) narrow down its solution space [19].

To this end, we view $\mathbf{x}$ as a random vector (equivalently each x_i as a random variable), and we derive additional constraints on the problem solution from the following statistical properties of $\mathbf{x}$: (i) lower and upper bounds of each component x_i of $\mathbf{x}$, (ii) lower and upper bounds of each component ratio x_i/x_j, and (iii) the prior probability distribution of $\mathbf{x}$. We refer to (i) and (ii) collectively as *range constraints*. The details are given next.

[4] While in theory one might expect matrix $\mathbf{C}$ to be rank deficient, in practice it generally is not, due to errors in the input measurements.

Introducing Range Constraints

We define $L_{i,i}$ and $U_{i,i}$ respectively as the lower and upper α percentiles[5] of the probability distribution of x_i. Similarly, we define $L_{i,j}$ and $U_{i,j}$ respectively as the lower and upper bounds of the quotient random variable $\rho_{ij} = x_i/x_j$, for any $i \neq j$. With a sufficiently small α, the α percentiles act as de facto lower and upper bounds of the respective random variable. If the values of $L_{i,j}$ and $U_{i,j}$ are known for all (i, j), the following set of inequality constraints apply:

$$L_{i,i} \leq x_i \leq U_{i,i} \qquad \text{for all } i \in \{1, 2, \cdots, n\} \tag{5}$$

$$L_{i,j} \leq \frac{x_i}{x_j} \leq U_{i,j} \qquad \text{for all } (i,j) \in \mathcal{I} \tag{6}$$

Thus, we can re-formulate the minimization problem in (3) as follows:

$$\min_{\mathbf{x}} \|\mathbf{C}\mathbf{x}\|^2 \text{ subject to } \begin{cases} L_{i,i} \leq x_i \leq U_{i,i}, \ 1 \leq i \leq n \\ L_{i,j} \leq \frac{x_i}{x_j} \leq U_{i,j}, \ (i,j) \in \mathcal{I} \end{cases} \tag{7}$$

It is easy to see that this is in effect a **linearly** constrained quadratic function minimization problem, hence a *quadratic programming* (QP) problem, since the inequalities in (6) can be expressed as linear inequalities in x_i and x_j:

$$x_i - U_{i,j} x_j \leq 0 \tag{8}$$

$$-x_i + L_{i,j} x_j \leq 0 \tag{9}$$

This quadratic program can be solved using standard (iterative) QP techniques; we currently use the `quadprog` function of the Matlab Optimization Toolbox. Also, a geometric interpretation of the solution is straightforward, as explained next.

Since a linear constraint defines a halfspace in n-dimensional space, the range constraints in (7) simultaneously define a convex bounded polytope in nD space. Let $\mathcal{P}$ denote this polytope, let $\mathcal{S}_0$ denote the one-dimensional unbounded solution space of our original (unconstrained) inverse problem (represented by Equation (4)), and let $\mathcal{S}_1 = \mathcal{P} \cap \mathcal{S}_0$. In the simple case $n = 2$, $\mathcal{S}_0$ is a line, $\mathcal{P}$ is a polygon, and their intersection consists of a line segment (or the empty set), as illustrated in Figure 2.

Now, if $\mathcal{S}_1 \neq \emptyset$, then $\mathcal{S}_1$ consists of a **bounded** one-dimensional subspace embedded in n-D space, and the solution of our QP is a point in $\mathcal{S}_1$ with the smallest residual $\|\mathbf{C}\mathbf{x}\|^2$. And if $\mathcal{S}_1 = \emptyset$, then the solution is a point in $\mathcal{P}$ (but not in $\mathcal{S}_0$) with the smallest residual $\|\mathbf{C}\mathbf{x}\|^2$. In other words, it is a point in $\mathcal{P}$ that is closest to $\mathcal{S}_0$.

[5] The lower α percentile of the probability distribution of a random variable X is a value z_L such that $\Pr(X \leq z_L) = \alpha$. Similarly, its upper α percentile is a value z_R such that $\Pr(X \leq z_R) = 1 - \alpha$.

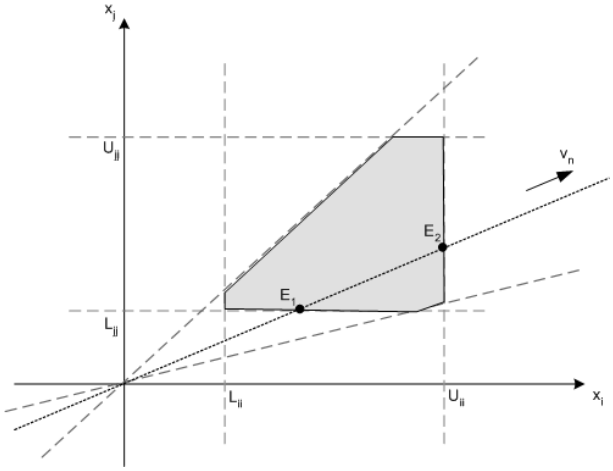


Fig. 2. Geometric interpretation of the solution space when $n = 2$, with and without range constraints-based regularization. Before regularization, the solution space consists of an unbounded one-dimensional subspace, $\mathbf{x} = k\mathbf{v}_n$ (*dotted line passing through origin*). The range constraints (*dashed lines*) define a convex polytope in nD space—a polygon when $n = 2$ (*gray-filled polygon*). Hence by imposing the range constraints, we obtain a solution that lies in the intersection of these two regions—-a bounded one-dimensional subspace (the *line segment between points* E_1 and E_2) .

Introducing Prior Probability

We extend the regularization method of Section 3.3 by using prior probability as an additional source of information in selecting a unique solution from the infinite set of candidates in (4). This is achieved by solving a sightly modified quadratic program:

$$\min_{\mathbf{x}}[\|\mathbf{C}\mathbf{x}\|^2 + \lambda\ell(\mathbf{x})] \text{ subject to } \begin{cases} L_{i,i} \le x_i \le U_{i,i},\ 1 \le i \le n \\ L_{i,j} \le \frac{x_i}{x_j} \le U_{i,j},\ (i,j) \in \mathcal{I} \end{cases} \tag{10}$$

where λ is a positive scalar parameter, and μ and $\mathbf{\Sigma}$ are respectively the mean vector and covariance matrix of the probability distribution of $\mathbf{x}$. Under the assumption that this distribution is Gaussian, the second term in the cost functional is proportional to the *Mahalanobis* distance of $\mathbf{x}$ from the population mean. And so this additional term has the effect of favoring candidate solutions with smaller Mahalanobis distance, hence higher prior probability, in the minimization process.

Obviously the geometric interpretation of the solution of this quadratic program remains the same as in Section 3.3, except that the solution is now a point that minimizes a different cost functional. It is also interesting to note that, assuming errors in the input measurements are of the form: $\mathbf{C}\mathbf{x} = \eta$,

where η is Gaussian white noise, then our solution of (7) in fact corresponds to the Maximum Likelihood estimate (MLE), and our solution of (10) corresponds to the Bayesian MAP (maximum a posteriori) estimate of our inverse problem.

3.4 Computing Length Ratios

Let us now re-visit the issue of how the length ratios (i.e. vector $\mathbf{y}$) are themselves computed from a single uncalibrated image. Basically, this is only possible under a limited set of conditions, each of which may require certain minimal calibration information. We use the following algorithm to determine what is needed for computing a length ratio (in addition to the projections of the line segments in the image) [5]:

- If the line segments are coplanar and their plane is almost parallel to the image plane, i.e. if the weak perspective camera model applies, then no further information is needed for the computation. (In this case, the length ratio of two 3D line segments is obtained as their length ratio in the image.)
- Otherwise, if the two line segments are collinear, then the vanishing point of their direction is needed for the computation.
- Otherwise, if the two line segments are parallel, then the vanishing point of their direction and the vanishing point of the line passing through an endpoint of each segment are needed for the computation.
- Otherwise, if the two line segments are coplanar, then the vanishing line of the plane containing them is needed for the computation.

4 Application

We now show that our novel visual metrology technique of Section 3 can be applied to the estimation of certain human-body anthropometrics from a full-body uncalibrated photograph of a person. Namely, the n line segments in the metrology technique consist of 10 anthropometrics (so $n = 10$), and each anthropometric is the straight-line distance between two visible and well-defined body landmarks (or canonical points). They are the following [20, 21]:

1. body height (or *stature*): between the base of the feet (or the standing surface) and the top of the head.
2. neck height: between the base of the feet and the trapezius—point at the side of the neck.
3. acromial height: between the base of the feet and the acromion—tip of the shoulder.
4. chin to top of head distance (or *head length*): between the top of the head and the chin (or menton) landmarks.

5. mouth to top of head distance: between the stomion—point at the center of the mouth—and top of head landmarks.
   item subnasale to top of head distance: between the subnasale—point at the base of the nose—and top of head landmarks.
6. forehead to chin distance: between the crinion—lowest point of the hairline on the forehead—and menton landmarks.
7. sellion to chin distance: between sellion—point at the deepest point of the nasal root depression—and the chin.
8. biocular distance: between outer corners of the eyes.
9. bitragion distance: between left and right tragions—points on the cartilaginous flaps in front of each earhole.

Clearly, for the first four anthropometrics the person is assumed to be in upright standing pose—not slouching or leaning on one side. Furthermore, assuming body pose and head pose are both (nearly) frontal with respect to the camera, these ten anthropometrics are all either (nearly) collinear or coplanar, and hence their ratios can be easily computed from an image (see Section 3.4). Specifically, anthropomerics 1–8 are all vertical and collinear, and so we only need the vertical vanishing point to compute their pairwise ratios. Anthropometrics 4–10 lie on the facial surface and are nearly coplanar, provided that the person is sufficiently far from the camera. With the head pose (nearly) frontal to the camera, the weak perspective model is a good approximation over the facial region (i.e. the facial anthropometrics are nearly coplanar in a plane nearly parallel to the image plane) and so we do not need any calibration information to compute the pairwise ratios of these facial anthropometrics. Currently we compute pairwise ratios of anthropomerics 1–8 and pairwise ratios of anthropomerics 4–10, hence a total of 39 ratios.

If, however, the whole body and/or the head are at an angle with respect to the camera, then the weak perspective model no longer applies in the facial region, and facial anthropometric ratios (4–10) cannot be easily estimated.

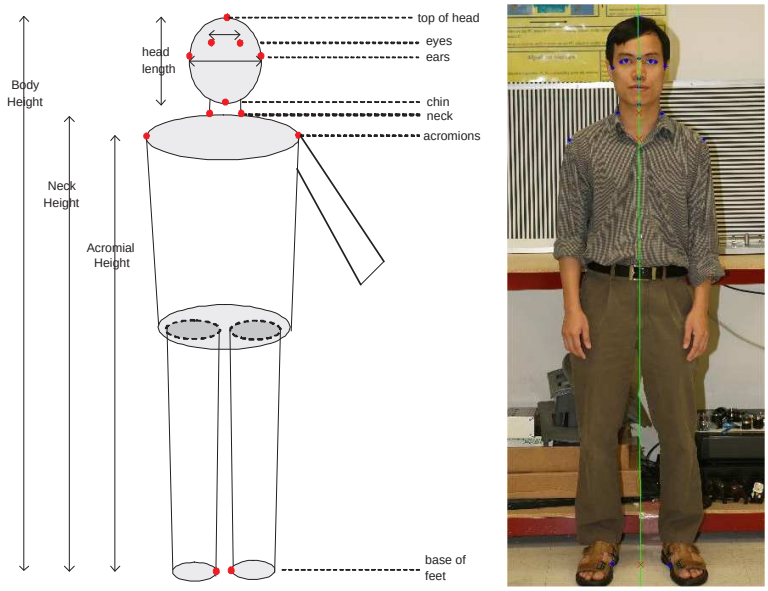
Finally, a word on the choice of anthropometrics. There are many other possible anthropometrics that can be used (in addition to or in place of the ones we chose to estimate in this paper). However, some factors have limited our choice, in particular: (i) ability to compute the anthropometric ratios, (ii) availability of anthropometric statistics, and (iii) accuracy of landmark localization in the image; for example, the bigonial width[6] is a bad choice because the gonial is difficult to locate in an image, even manually and even with high-resolution images.

4.1 Landmark Localization

Localization of the landmarks associated with the ten anthropometrics of interest is achieved as follows; first we locate the following 13 body landmarks

[6] Bigonial (or mandible) width: the straight-line distance between the left and right *gonial* landmarks—the corners of the jaw.

in the image: top of the head, forehead, subnasale, stomion, chin, left and right outer corners of the eyes, left and right tragions, left and right trapezius, left and right neck points, left and right acromions, and left and right medial longitudinal foot arches. This is currently done semi-automatically by having the user select points in the image via an interactive Matlab interface. We then estimate the person's medial axis (or midline of symmetry) as the line passing through the following four points: top of the head, chin, midpoint of the two foot landmarks, and the vertical vanishing point. We refine the locations of the first five landmarks (top of head, forehead, subnasale, stomion, chin) by projecting them onto the medial axis. For the landmark pairs of the trapezius, neck, acromions, and feet, we compute the intersection point of the medial axis with the line segment joining each landmark pair. This way, the vertical anthropometrics are all collinear with the medial axis. See Figure 3(b).



(a) (b)

Fig. 3. (a) Some of the anthropometrics we estimate in this paper, and (b) the body landmarks associated with them; the *blue dots* are points we mark manually; the *green line* is medial axis of symmetry of the person, which we estimate by fitting a line to some of these points (see text).

It is important to note that landmark localization is prone to error even when done manually. The main sources of error include: non-frontal body pose, occlusion by body hair and/or clothing, and fuzziness of the landmark.

4.2 Anthropometric Statistics

The statistics required for our anthropometric estimation method are obtained (directly or derived) from the seminal anthropometric survey in [22], and they are all categorized by gender, i.e. separate statistics for males and females. Table 1 shows these statistics for some anthropometrics based on α =1e-12. The details of how they were all obtained are given next.

- The mean and standard deviation of each anthropometric, and the correlation coefficient of each anthropometric pair are taken from the survey.
- We compute the upper and lower bounds of each anthropometric as the upper and lower 100α percentiles of the normal distribution $N(\mu, \sigma)$, where μ and σ are respectively its mean and standard deviation.
- The covariance matrix, $\mathbf{\Sigma}$, of the anthropometrics is computed from their standard deviations and pairwise correlations. Namely, if σ_i and σ_j are the standard deviations of the ith and jth anthropometrics, respectively, and ρ_{ij} is their correlation coefficient, then: $\mathbf{\Sigma}_{ij} = \rho_{ij}\sigma_i\sigma_j$.
- The upper and lower bounds of each anthropometric ratio are computed from the means and standard deviations of the two associated anthropometrics. We first estimate the probability distribution of the anthropometric ratio using the formula in [23], then obtain the upper and lower bounds respectively as the upper and lower 100α percentiles of this distribution.

Table 1. Some anthropometric statistics (in cm), where SD: standard deviation, UB: upper bound, LB: lower bound. Based on the 1988 US Army survey [22] and α =1e-12 (see text).

	Males				*Females*			
	Mean	SD	LB	UB	Mean	SD	LB	UB
Body height	175.5	6.68	128.8	222.3	162.9	6.36	118.4	207.5
Head length	23.20	.88	17.04	29.36	21.76	.85	15.81	27.71
Bitragion	14.48	.60	10.28	18.68	13.64	.52	10.0	17.28
Interocular	10.2	.54	6.42	13.98	9.62	.50	6.12	13.12

5 Experiments and Results

5.1 Data and Methodology

We tested our anthropometric estimation method using both synthetic data and real images. The former consists of randomly generated $\mathbf{y}$ vectors, each of which is obtained by first randomly generating a vector $\mathbf{x}$ from the multivariate Gaussian distribution of male or female anthropometrics (Section 4.2), then computing the corresponding pairwise ratios. See Figure 4.

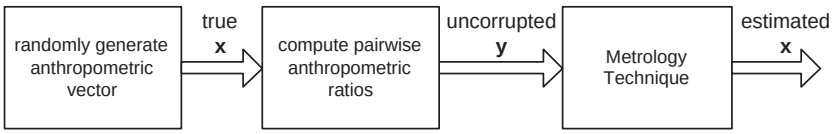


Fig. 4. Methodology for (random) generation of synthetic anthropometric data (namely **y** vectors) to be used in testing our anthropometric estimation method.

The real images consist of a set of high-resolution (4368x2912) images captured in-house using a Canon 28-200mm EOS camera (Fig. 5). Each image is a **full-body frontal shot** of one person. The dataset contains a total of 108, of 27 different adults, 7 females and 20 males, spanning various ethnicities: caucasian, chinese, indian, and african. The vertical vanishing point needed to compute pairwise ratios of anthropometrics 1–8 is computed as the intersection in the image of parallel vertical lines of the background scene.

Fig. 5. Sample images from our inhouse dataset.

Because all our images are frontal shots with the person sufficiently far from the camera, we only need the vertical vanishing point to compute the pairwise anthropometric ratios (Section 4). This vanishing point is estimated as the intersection in each image of at least two vertical lines of the background scene (which are available in all our images). The lines are in turn obtained by manually marking five points in the image for each vertical line, and fitting them to a line via least squares estimation. We should also note that our method assumes the gender of the person in the image is known, and uses the anthropometric statistics of the corresponding gender (male or female). The method is applied with the following parameter values: α =1e-7, and λ =.01. The estimation error is defined as: true value - computed value, and the results are presented in the following subsections.

5.2 Synthetic Data Results

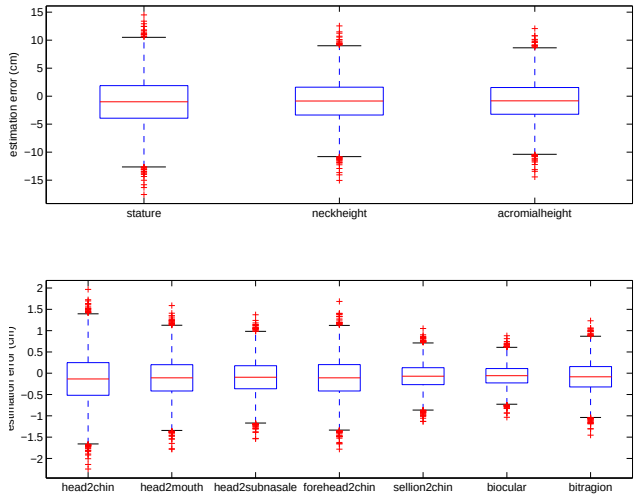
These results are summarized in Figure 6 based on **5000** random vectors divided equally between males and females. Notice that since the synthetic data vectors are uncorrupted by noise (apart from computer roundoff errors), the estimation error in this case is equal to the method's *model error*. In other words, it does not reflect any kind of input error. Plot (a) shows the estimation error distribution as a boxplot [7]. For example, the standard error is **4.3cm** for body height and **.58cm** for head length.

The plot in (b) shows the estimation error for body height versus the true anthropometric value. Clearly, it suggests that estimation error is somewhat negatively correlated with true body height—the computed correlation coefficient is -.46. This means estimation error tends to be larger for people with "extreme" body height, i.e. who are far from the population mean. Plot (c) shows the same thing as the plot in (b) but for head length, again suggesting a negative correlation between head length and its estimation error—the computed correlation coefficient is -.51. This result should not be too surprising, since our metrology technique favors the candidate solutions with larger prior probability, hence close to the population mean (Section 3.3). The plot in (d) shows the distributions of body height estimation separately for males and females. A one-way analysis of variance has rejected the hypothesis that these two distributions are significantly different.

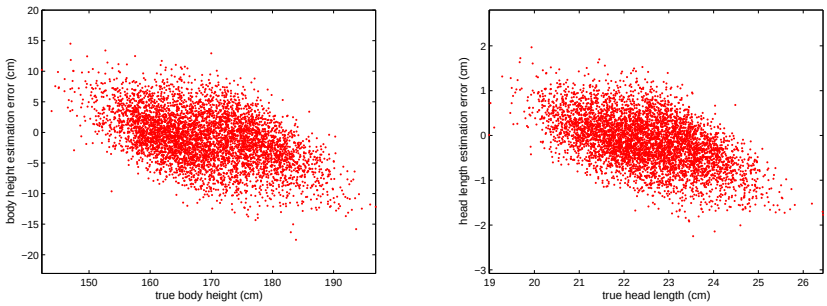
5.3 Real Images Results

These results are summarized in Figure 7, but only pertaining to body height estimation error, because (unfortunately) we do not have groundtruth values for all other anthropometrics. Plot (a) shows the estimation error distribution (again as a boxplot), which has a mean of **-7.1cm** and a standard deviation of **6.1cm**. As expected, estimation error is significantly larger for real images than synthetic data, since this error reflects both model and input errors. Plot (b) shows the error distributions separately for males and females. A one-way analysis of variance has rejected the hypothesis that these two distributions are significantly different. However, visually speaking the female distribution appears to have a smaller range of variation, which may be an artifact of the smaller female sample in our dataset—20 males vs. 7 females. Plot (c) shows that as with synthetic data estimation error is negatively correlated with true body height, and indeed the computed correlation coefficient is -.49. In particular, the largest errors are obtained for the shortest and tallest persons in the dataset.

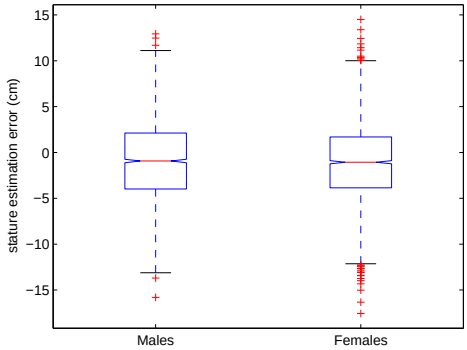
[7] A boxplot is a way of displaying the distribution of a data sample. At the center is a box with lines at the lower quartile, median, and upper quartile values. The two lines extending from each end of the box (called *whiskers*) represent the middle 75% of the data. The '+' marks beyond the ends of the whiskers are called *outliers*.



(a)

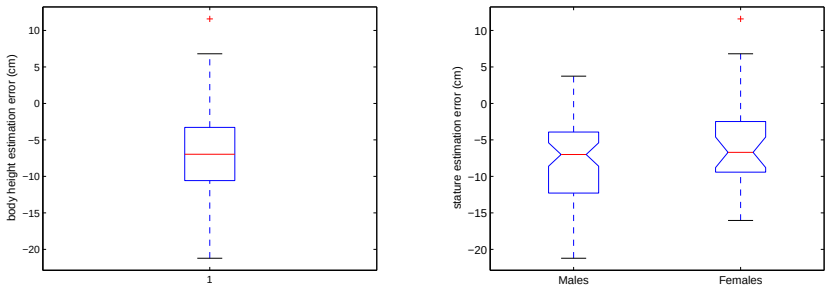


(b) (c)

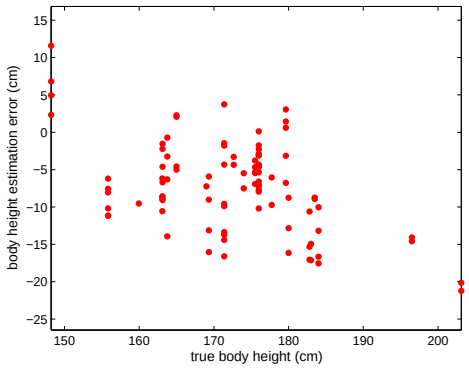


(d)

Fig. 6. Estimation results based on 5000 synthetic **y** vectors: (a) distribution of estimation error for all 10 anthropometrics; (b) true body height versus estimation error; (c) true head length versus estimation error; (d) male and female distributions of body height estimation error.



(a) (b)



(c)

Fig. 7. Estimation results based on real images: (a) distribution of body height estimation error; (b) male and female distributions of body height estimation error; (c) true body height versus estimation error.

6 Conclusions and Future Work

We have developed a novel method for estimating anthropometrics from a single un-calibrated image. The method may need minimal calibration information for computing anthropometric ratios from the image (basically vanishing points and/or line), but requires *no* absolute length information (i.e. a reference length). In a nutshell, we compensate for this information via clever use of prior knowledge about anthropometric statistics. Naturally, the estimation accuracy is inferior to other methods that use absolute length information and/or more cameras, such as [4, 1]. However, our goal is not to unconditionally achieve very accurate estimation, but rather to obtain as accurate an estimation as possible with limited input information. Besides, inaccurate

estimates can be useful when combined with other (inaccurate) knowledge sources. The method is primarily targeted for forensic-style identification, but can also be useful for tracking people across multiple camera views in surveillance applications, provided that the landmark localization and gender classification parts are automated. Ongoing and future work includes:

- Automate localization of body landmarks.
- Study the effect on estimation accuracy of the control parameter values of the method (i.e. α and λ).
- Study the effect on estimation accuracy of image resolution and head pose variation.
- Extend the method to non-frontal body poses.

Acknowledgments

This work has been supported in part by the US Office of Navy Research (ONR), grant N000140610102.

References

1. P. Meunier and S. Yin, "Performance of a 2d image-based anthropometric measurement and clothing sizing system," *Applied Ergonomics*, vol. 31, no. 1, pp. 445–451, 2000.
2. N. Saitoh, K. Kurosawa, and K. Kuroki, "A study on height measurement from a single view," in *IEEE International Conference on Image Processing*, 1999.
3. A. Criminisi, A. Zisserman, L. Van Gool, S. Bramble, and D. Compton, "A new approach to obtain height measurements from video," in *Proc. of SPIE*, 1998, vol. 3576.
4. A. Criminisi, I. Reid, and A. Zisserman, "Single view metrology," *International Journal of Computer Vision*, vol. 40, no. 2, pp. 123–148, Nov. 2000.
5. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, second edition, 2004.
6. Feng Guo, *Calibration and Metrology Using Still and Video Images*, Ph.D. thesis, University of Maryland, College Park, 2007.
7. X. Cao and H. Foroosh, "Metrology from vertical objects," in *British Machine Vision Conference*, 2004.
8. C. Madden and M. Piccardi, "Height measurement as a session-based biometric," in *Image and Vision Computing New Zealand*, 2005.
9. A. Bovyrin and K. Rodyushkin, "Human height prediction and roads estimation," in *IEEE Conference on Advanced Video and Signal-Based Surveillance*, 2005.
10. Biswajit Bose and Eric Grimson, "Ground plane rectification by tracking moving objects," in *IEEE Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, 2003.

11. Jochen Meidow, "Calibration of stationary cameras by observing objects of equal heights on a ground plane," in *Proc. ISPRS Congress*, 2004, pp. 1067–1072.
12. F. Lv, T. Zhao, and R. Nevatia, "Camera calibration from video of a walking human," vol. 28, no. 9, pp. 1513–1518, 2006.
13. Xiaofeng Ren, Alexander C. Berg, and Jitendra Malik, "Recovering human body configurations using pairwise constraints between parts," in *ICCV*, 2005, pp. 824–831.
14. I. Mikic, M. M. Trivedi, E. Hunter, and P. C. Cosman, "Human body model acquisition and tracking using voxel data," *International Journal of Computer Vision*, vol. 53, no. 3, pp. 199–223, 2003.
15. Alexei Gritai, Yaser Sheikh, and Mubarak Shah, "On the use of anthropometry in the invariant analysis of human actions," in *International Conference on Pattern Recognition*, 2004, p. 923.
16. C. Barron and I. Kakadiaris, "Estimating anthropometry and pose from a single uncalibrated image," *Computer Vision and Image Understanding*, vol. 81, pp. 269–284, 2001.
17. C. Barron and I. Kakadiaris, "On the improvement of anthropometry and pose estimation from a single uncalibrated image," *Machine Vision and Applications*, vol. 14, pp. 229–236, 2003.
18. Gilbert Strang, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, third edition, 2003.
19. Curtis R. Vogel, *Computational Methods for Inverse Problems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002.
20. Leslie G. Farkas, Ed., *Anthropometry of the Head and Face*, Raven Press, second edition, 1994.
21. Stephen Pheasant, *Bodyspace: Anthropometry, Ergonomics and the Design of Work*, Taylor and Francis, second edition, 1996.
22. J. Cheverud, C. G. Gordon, R. Walker, C. Jacquish, L. Kohn, A. Moore, and N. Yamashita, "1988 anthropometric survey of us army personnel," Tech. Rep. TR-90/032, US Army Natick Research Labs, May 1990.
23. D. V. Hinkley, "On the ratio of two correlated normal random variables," *Biometrika*, vol. 56, no. 3, pp. 635–639, 1969.