The Reliability of Forensic Body-Shape Identification

Neerja Thakkar, Georgios Pavlakos, Hany Farid
Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley CA, USA
nthakkar@berkeley.edu, pavlakos@berkeley.edu, hfarid@berkeley.edu

Abstract

Photo-based forensic identification can be critical in the prosecution of, and defense against, criminal charges. Identification techniques range from the specific biometric-based to the more generic, based on height, weight, gender, and race. Although fairly basic, accurate height and weight estimation remains challenging due to physiological factors, concealing clothing, body pose, and the scale ambiguity inherent to the photographic process. We describe an extension to 3D body-pose estimation that more accurately estimates body shape across a broader range of body sizes. We evaluate the reliability of this technique in making metric estimates of height and weight, and in making non-metric categorization of people based on a scale-agnostic measure of body shape. Although this approach improves on previous efforts, we find that accurate body-shape identification from a single, reference-free image remains challenging.

1. Introduction

On January 6, 2021 a mob of some $2,000$ attacked the Capitol Building in Washington D.C. seeking to disrupt the congressional confirmation of President-elect Joe Biden's electoral victory. In the aftermath, five lay dead and hundreds were injured. In the weeks and months that followed, federal and state law enforcement asked the public for help in identifying those involved in the assault, including an FBI tipline (https://www.fbi.gov/uscapitol1) asking members of the public to submit photos and videos of the violence. In some cases, those involved were readily recognizable from their unmasked faces in the photos and videos. In other cases, however, those involved were masked or obscured in a larger crowd, Figure 1.

With a camera in most pockets, the average citizen is increasingly likely to document everything from street-level crime to large-scale protests and riots. At the same time, photographic identification from images – beyond well-studied techniques for face recognition [21] – can still pose significant challenges. Determining something as basic as a person's height and weight, for example, is riddled with complexity due to the inherent loss of information resulting from the 3D to 2D image projection and the resulting scale ambiguity. Additional complicating factors arise from natural daily fluctuations in height and weight: spinal compression results in a daily change in height by as much as 2 cm [16], and body weight similarly fluctuates daily by as much as ± 2.25 kg in a single day [5]. Differences in body pose add further complications: slouching or walking, for example, can result in an apparent change in height by as much as 6 cm [11].

Figure 1. How reliably can identifying information be extracted from this photo of masked Capitol Hill rioters? [source: Getty Images]

While height and weight alone are not sufficient to uniquely identify a person, these basic measurements (along with race, gender, and age) can be useful in narrowing the field of potential suspects, or eliminating a suspect from consideration.

Building on recent advances in 3D human modeling [20, 24, 26, 30], recent work [31] investigated if 3D human modeling can reduce some of the ambiguities and uncertainty in forensic height and weight estimation from a single, reference-free image. This work showed that by ad-

justing for 3D pose, height estimation can be significantly improved. At the same time, however, 3D modeling did not improve weight estimation beyond guessing a gender-specific average weight. The authors cite two primary limitations to height and weight estimation: (1) for height and weight estimation, resolving the scale ambiguity in the absence of known reference objects in the scene; and (2) for weight estimation, the human 3D modeling is not able to estimate the shape of below or above average-sized bodies, even while accurately estimating body pose. In our work, we seek to address these two limitations to determine if more accurate 3D modeling will yield a sufficiently accurate tool for forensic identification.

In the first part of our work (Section 3), we augment the 3D modeling of [30] to more accurately estimate body shape across a range of body sizes, and show how this augmented modeling improves height and weight identification (Section 4). Despite these improvements, we find that resolving the scale ambiguity from a single, reference-free image remains a significant challenge for accurate forensic identification.

In the second part of our work (Section 5), we propose a scale-agnostic classification of body shape as an alternative to the classic measurements of height and weight. In particular, we evaluate the reliability and distinctiveness of the estimated low-dimensional, body-shape parameterization for use in forensic identification.

2. Related Work

2.1. Height and Weight

Classic approaches to height estimation require the person in question to be standing upright and next to an object of known size [10, 11]. These approaches, however, are unable to contend with different body poses or reference-free scenes that often arise in forensic settings.

Among different approaches to height estimation [2, 8, 25, 29], BenAbdelkader et al. [4] proposed a reference-free approach to estimating human height that combines classic single-view metrology with statistical knowledge of human anatomy. The accuracy of this approach, however, is only slightly better than guessing a gender-specific average height [31]. Similarly, Zhu et al. [34] use priors learned by a neural network to recover the absolute scale of a scene from a single image. This approach is worse than guessing a gender-specific average height [31]. And, Bieler et al. [6] use explicit knowledge of gravity to measure a person's height from a video sequence. While more accurate than the two previous approaches, this approach is limited to video in which the person being measured is in free fall.

As compared to height estimation, there are fewer approaches to estimating body weight. Velardo and Dugelay [32] estimate weight from manually-extracted body measurements extracted from a frontal and side view image containing a reference object, achieving a mean average error (MAE) of 7.2 kg. By comparison, Arigbabu et al. [3] estimate weight by training a neural network with 13 measurements of the human body across multiple video frames, yielding a MAE of 6.4 kg. And, Nguyen et al. [28] estimate weight from a single color and depth image, yielding a MAE of 4.6 kg. These approaches perform better than guessing a gender-specific average weight, which would yield an average error of 15.8 kg. In contrast to these approaches, our method requires only a single image and does not assume the presence of a reference object.

Figure 2. An input image and manually extracted binary silhouette (top row). A fitted 3D SMPLX model without (bottom left) and with (bottom right) the additional silhouette-based, body-shape constraint. Although body pose is well estimated in both cases, the additional constraint yields a more accurate estimate of body shape. [photo source: Official White House Photo by Shealah Craighead]

2.2. Body Shape

Prior approaches have considered the use of anthropometric measurements as a biometric. Godil et al. [17] explored the use of anthropometric measurements, including, for example, the 3D lengths of the wrist to elbow, elbow to shoulder, and hip to knee. Other works have investigated the feasibility of using the shape of specific body parts as a biometric, including hand shape [14], ear shape [33], lip shape [18], and eyebrow shape [13]. These techniques are concerned with identifying individuals from distinct biometric features. In contrast, we focus on a more generic forensic identification based on height, weight, and overall body shape.

A few approaches have explored the use of body-shape identification. Using millimeter wave imaging, Gonzalez et al. [19] extract body shape in a fixed neutral position, from which they characterize the 2D shape of the human body. The benefit of this approach is that the millimeter wave imaging penetrates any obscuring clothing. In related work, Kalyanaraman et al. [22] extract body shape from a radar sensor for the purposes of in-home soft biometrics (i.e., distinguishing someone from a relatively small sample size). The drawback of both of these approaches is they require specialized imaging hardware and operate in 2D. In contrast, we extract 3D body models from a single, standard RGB image.

3. Human 3D modeling

3.1. SMPLify-X

We employ SMPLify-X [30] to estimate 3D body pose from a single image. This model fitting approach uses a 3D SMPLX model of the human body, $M(\theta, \beta, \psi)$, parameterized by body pose, θ, body shape, β, and facial expression ψ. The SMPLX model extends the popular SMPL model [26], allowing for articulated hands and more accurate and detailed facial expression modeling. A specified body shape β can be rendered in one of three user-selected styles: feminine, masculine, or neutral, each of which adopts somewhat conventional, gender-based body shapes. Importantly, the same β representation rendered under different styles will yield different body shapes. That is, the β parameters alone do not uniquely define body shape.

This 3D model fitting consists of two basic steps: (1) from a single image, use OpenPose [7] to detect 2D keypoints from the body, face, hand, and feet; and (2) from the 2D keypoints, simultaneously estimate the 3D body parameters θ, β, ψ and the camera extrinsic parameters, by minimizing the difference between the detected 2D keypoints and the 3D keypoints reprojected into the image plane [30].

Importantly, the 3D model estimation does not consider the original shape of the body in the image, and incorporates several priors on human-body shape. This model estimation, therefore, struggles to estimate the body shape of below or above average-sized bodies, even while estimating accurate body pose, Figure 2 (bottom left). As a result, weight estimation is severely hampered as compared to height estimation [31].

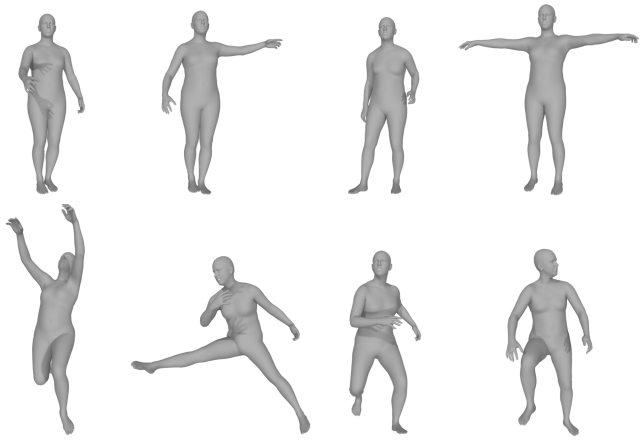

Figure 3. Examples from the simulated data set of neutral (top) and action (bottom) poses [31].

3.2. SMPLify-X with Improved Body Shape

The original SMPLify-X pipeline [30] optimizes the following objective function:

$$\begin{aligned} E(\beta, \theta, \psi) &= E_J + \gamma_{\theta_b} E_{\theta_b} + \gamma_{m_h} E_{m_h} + \gamma_{\theta_f} E_{\theta_f} \\ &+ \gamma_\alpha E_\alpha + \gamma_\beta E_\beta + \gamma_C E_C + \gamma_\psi E_\psi, \quad (1) \end{aligned}$$

where E_J is a joint re-projection loss, E_{θ_b} is a variational-autoencoder (VAE) based body-pose prior, and where the other terms represent L_2 penalties on hand pose (E_{m_h}), facial pose (E_{θ_f}), extreme bending of elbows and knees (E_α), body shape (E_β), interpenetration (E_C), and facial expression (E_ψ), each with hand-tuned scalar weights γ_*.

We next describe an extension of the SMPLify-X 3D body estimation that incorporates explicit knowledge of the 2D body silhouette in the original image. Because even modern techniques are not able to reliably perform figure-ground segmentation, and because in our forensic application it is permissible to have a human in the loop, we allow for an analyst to manually segment the person of interest from the background. From the original image, we begin by manually extracting a binary-valued mask segmenting the entire body from the background, Figure 2. On each step of the iterative optimization, a differentiable renderer [23] is used to generate a binary-valued mask of the current version of the 3D model reprojected into the 2D image plane.

The objective function, Equation (1) is augmented with the L_2 norm of the difference between the silhouette (Figure 2) and the 2D reprojection of the 3D model. Specifically, to the original objective function, a new shape term,

$\gamma_S E_S$, is added, where $\gamma_S = 3000$. In addition, the shape prior weight γ_β is reduced by a factor of 10, to allow for more deviation from the prior and increased reliance on the silhouette.

Shown in Figure 2 is an example of the 3D body modeling without (bottom left) and with (bottom right) the additional constraint. Here we see that the additional constraint leads to more accurate body-shape modeling.

4. Height and Weight Identification

4.1. Data Set

In order to evaluate the accuracy of height and weight estimation, we create a simulated dataset of 3D SMPLX models. This dataset allows us to evaluate our method across a variety of body shapes, camera angles, and poses. This dataset is generated by first sampling three body shapes of neutral gender from the SMPLX β-space, corresponding to a small, medium, and large body shape. These models are then posed into one of 12 different poses, ranging from neutral to action poses, Figure 3. Each shape in each pose is rendered with a virtual camera in one of two elevations (0 or 18 degrees) and three azimuths (-30, 0, or 30 degrees), yielding a total of 216, 800×800 pixel images.

4.2. Measurement

Each of the 3D models described above is subjected to the original and augmented body-fitting procedure described in the previous section. Then, starting with the estimated body pose θ – specified as a 32D latent vector – VPoser [30] is used to first repose the estimated 3D model into a neutral standing pose. Once reposed, height is measured as the distance from the top of the head to the plane formed by three points on the bottom of the feet. The interior volume of a 3D model is measured as a proxy for weight.

While these height and weight calculations are straightforward, they require that the 3D model be specified in real-world units. Although the fitted 3D SMPLX model is estimated in real-world units, this metric reconstruction is not accurate enough to support accurate height and weight estimation [31]. We, therefore, adopt two alternate approaches to estimating absolute scale as described in [31].

First, because the adult inter-pupillary distance (IPD) – measured as the distance between the center of the two pupils – is relatively similar for women and men [12], either the average IPD or a known person-specific IPD can be used to scale the full 3D model. Second, in order to disentangle the impact of the underlying 3D pose and shape estimation from the scale ambiguity, we align the estimated 3D model to the ground-truth 3D model using coherent point drift (CPD) [27]. CPD is a point-set registration algorithm that estimates the 3D rotation, isotropic scaling, and translation between two arbitrary point clouds (in our case, the point clouds correspond to the vertices of the underlying 3D models). After alignment, the height and weight are estimated in the units of the ground-truth model. This CPD-based approach is, of course, only applicable in simulation and not in real-world scenarios.

Table 1. Mean (variance) height and weight estimation errors reported as percent error relative to ground truth, corresponding to CPD- or IPD-based scale disambiguation, without ($\gamma_S = 0$) or with ($\gamma_S > 0$) the additional body-shape constraint.

condition	height	weight
$\gamma_S = 0$, CPD	0.75 (0.50)	21.28 (157.00)
$\gamma_S = 0$, IPD	5.01 (28.52)	33.51 (294.47)
$\gamma_S > 0$, CPD	0.78 (0.69)	8.07 (41.34)
$\gamma_S > 0$, IPD	3.24 (7.90)	17.45 (124.3)

4.3. Evaluation

We report the accuracy of height and weight estimation without (Section 3.1) and with (Section 3.2) the extension of the 3D body modeling incorporating overall body shape.

Shown in Table 1 are the mean errors in height and weight estimation averaged over all camera angles and body poses. For the full-reference CPD scale disambiguation, the addition of the body-shape constraint ($\gamma > 0$) has little impact on height estimation, but leads to significant improvement in weight estimation with a reduction in mean error from 21.28% to 8.07%. For the IPD scale disambiguation, the addition of the body-shape constraint leads to a slight improvement in height estimation with a reduction in mean error from 5.01% to 3.24%, and a significant improvement in weight estimation with a reduction in mean error from 33.51% to 17.45%.

In both the case with ($\gamma > 0$) and without ($\gamma = 0$) the additional body-shape constraint, estimation for neutral poses is more accurate than for action poses. For $\gamma > 0$ and CPD-based scale disambiguation, the mean height/weight error is 0.48%/6.86% for neutral poses compared to 1.21%/9.75% for action poses; and for IPD-based scale disambiguation the mean height/weight error is 2.63%/14.85% for neutral poses compared to 4.11%/21.07% for action poses.

Average U.S. adult female/male heights are normally distributed with a mean of 161/175 cm and a standard deviation of 7.0/7.4 cm [1]. Similarly, average U.S. adult female/male weights are normally distributed with a mean of 78.7/90.8 kg and a standard deviation of 19.7/19.8 kg [1]. If a person's height is estimated to be the average, gender-specific height, the average height estimation error for women/men would be 3.5%/3.2%. And, if a person's weight is estimated to be the average, gender-specific weight, the average weight estimation error for women/men

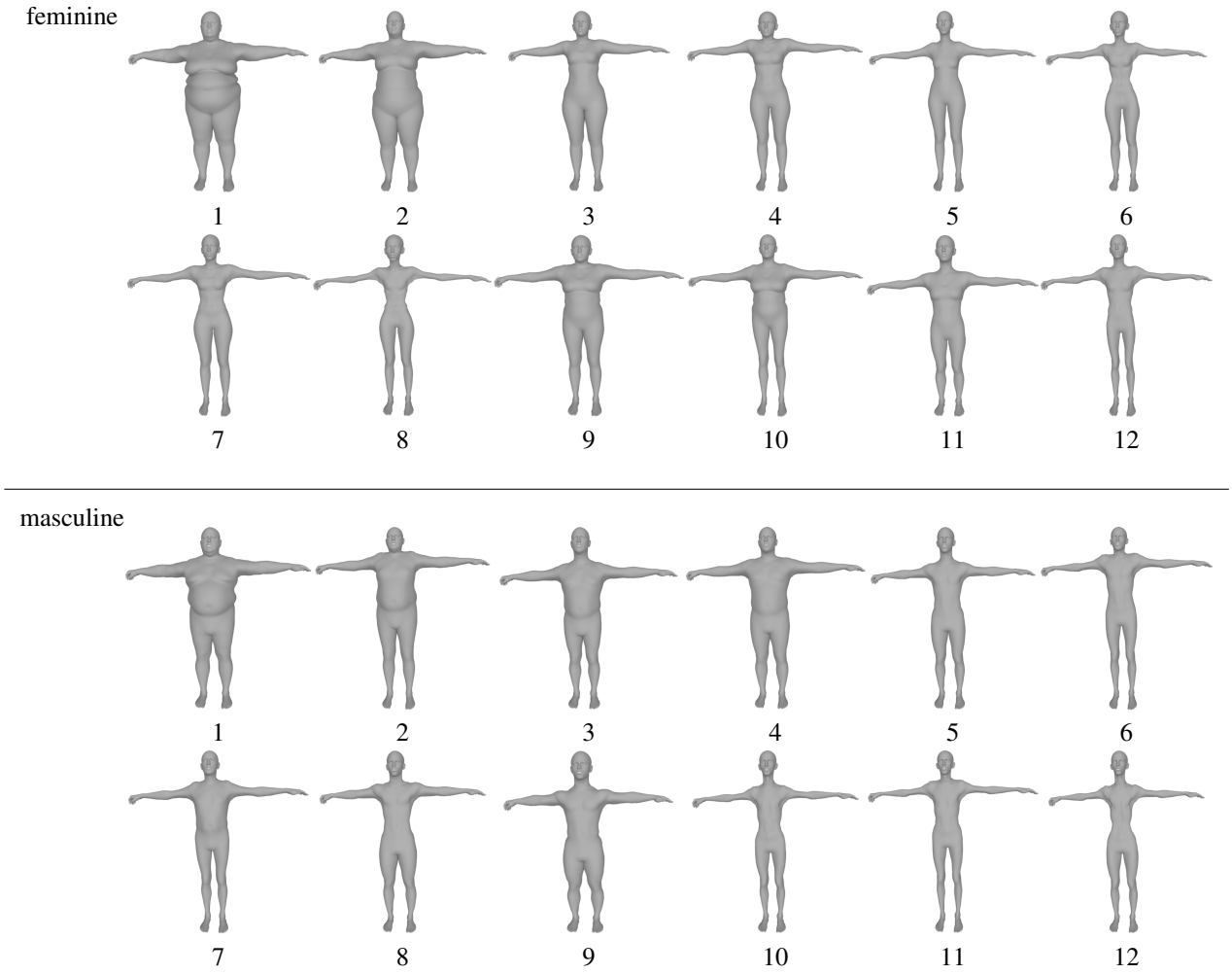

Figure 4. Representative views of the 12 feminine (top) and 12 masculine (bottom) equivalence classes.

would be $20.0\%/17.3\%$.

By comparison to this baseline, 3D body pose and shape estimation along with a perfect estimate of scale (CPD) leads to reasonably accurate estimates of height and weight. For purposes of identification, however, reliable estimation of height and weight remains challenging. Assuming, for example, a person of average height and weight of 175 cm (5', 9") and 90 kg (198 lb), with a 95% confidence[1], this person's height can be estimated to within a range of $170-177$ cm (5', 7" - 5', 10") and $69-103$ kg ($152-227$ lb). Note that the height and weight are consistently under reported due to a slight underestimation in body shape estimation.

With a less precise estimate of scale (IPD), estimates of height and weight are similar to guessing a gender-specific average. These estimates, of course, could be made more accurate with better estimates of scale. An accurate estimate of absolute scale, however, is notoriously difficult, particularly from a single reference-free image in uncontrolled environments. We, therefore, next propose a scale-agnostic technique for classifying human shape and size, the details of which are described next, along with a comparison of identification distinctiveness relative to the classic height and weight estimation.

[1]Because the height and weight errors are normally distributed with mean μ and standard deviation σ, the 95% confidence interval is specified as $\mu \pm 2\sigma$.

5. Body-Shape Identification

5.1. Equivalence Classes

A data-driven approach is employed to automatically discover the range of distinct body shapes – binned into discrete equivalence classes – spanned by the SMPLX β-space. The SMPLX model parameterizes body shape as a 10D, real-valued vector $\vec{\beta}$. In order to span a wide range of shapes while avoiding an excessive number of monster-sized or emaciated bodies, each component of a 10D vec-

tor, $\vec{\beta}$, is drawn randomly from the range $[-3, 3]$, with a randomly selected component of $\vec{\beta}$ drawn from the range $[-5, 5]$. A total of $50,000$ body shapes are randomly generated, and the feminine, 3D model associated with each of these body shapes is constructed in a neutral pose (recall that SMPLX models are rendered as feminine, masculine, or neutral). In order to compare body shape independent of scale, each rendered body is aligned to a single reference model using CPD.

These $50,000$ feminine body shapes are then iteratively partitioned into a discrete number of equivalence classes. To begin, a randomly selected body shape $\vec{\beta}_1$ is assumed to be the representative body shape for the first equivalence class C_1. Next, the similarity between the representative body for C_1 and the next body shape $\vec{\beta}_2$ is computed using the measure of intersection-over-union (IoU) applied to the binary silhouette of the rendered body shapes in the same, aligned, and neutral pose. If the IoU (with a maximum value of 1 corresponding to identical shapes) is greater than a specified threshold of 0.85, then $\vec{\beta}_2$ joins class C_1, otherwise $\vec{\beta}_2$ is distinct from $\vec{\beta}_1$ and becomes the representative shape for a new equivalence class C_2. This process is repeated with each new body shape $\vec{\beta}_i$ being compared to the representative body shape for each equivalence class, and either added to an existing class (with the maximum IoU above 0.85) or used to seed a new class (if the maximum IoU is below 0.85).

This iterative process yields a total of 19 feminine equivalence classes from the initial $50,000$ body shapes. This entire process was repeated with a new set of $50,000$ body shapes rendered with a masculine style, yielding a total of 15 equivalence classes (working within the gender constraints of the SMPLX model, we adopt this simplistic binary, gender categorization).

These 19 feminine, and 15 masculine equivalence classes were then manually pruned to 12 each by removing clearly emaciated or unrealistic body shapes, and (because IoU is not a perfect measure of body similarity) collapsing perceptually similar classes. Shown in Figure 4 and 4 are the representative body shapes for the final feminine (top) and masculine (bottom) equivalence classes.

5.2. Distinctiveness

The distinctiveness of the feminine and masculine equivalence classes is evaluated by building a pair of 12-way logistic regression models. The process in the previous section yielded a minimum of 1000 body shapes per equivalence class (any class with more than 1000 shapes was randomly pruned down to 1000 shapes). These shapes were randomly split into a $80/20$ training/testing dataset, from which a logistic regression model was trained to predict the gender-specific, body class from the underlying body-shape beta parameterization.

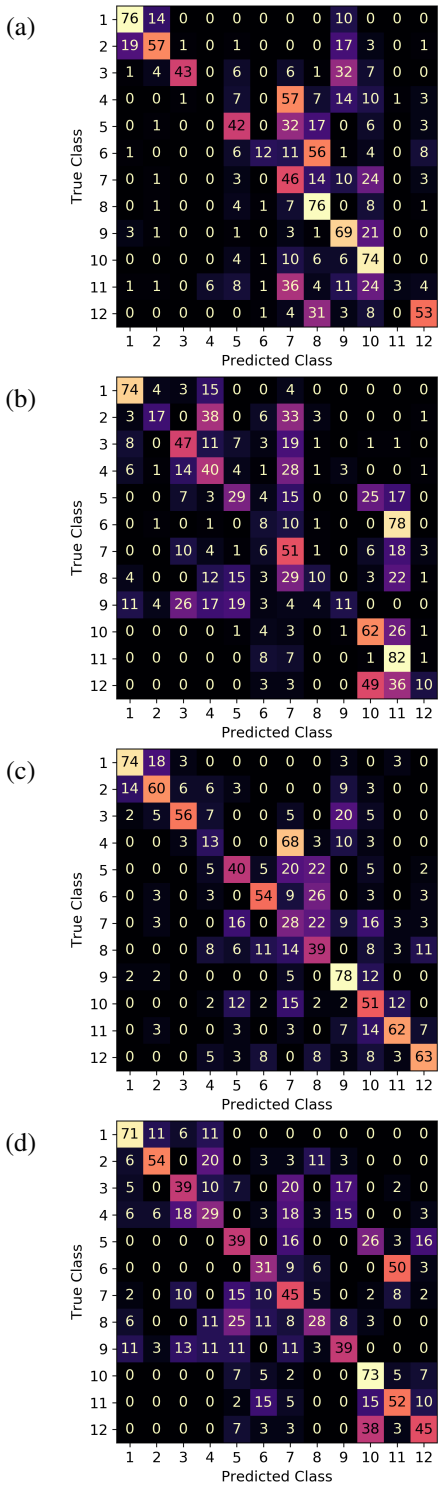

Figure 5. Confusion matrices evaluating body shape distinctiveness for (a,c) feminine and (b,d) masculine, for the (a,b) original and (c,d) augmented classifier. Values are reported as percentages (values less than 0.5% are rounded down to 0). Each class label refers to one of 12 body shapes shown in Figure 4.

The testing accuracy of these classifiers is 94.1% (feminine) and 95.0% (masculine). We see that although some of the shapes in Figures 4 are perceptually similar, their underlying beta parameterizations are distinct, suggesting that these shapes can be used for forensic identification. In order to be applicable in a general forensic setting, however, we must determine if the body-shape parameters can be reliably estimated from images depicting a range of different body poses and camera angles.

The representative body shape for each equivalence class was rendered in one of 12 neutral and action poses, Figure 3, and with the virtual camera in one of two elevations (0 or 18 degrees) and three azimuths (-30, 0, or 30 degrees). This yielded a total of $864,800 \times 800$ pixel images per gender.

As described in Section 3.2, a SMPLX model was fit to each image to yield an estimated body shape $\vec{\beta}$. Each of these body-shape parameterizations (with, of course, known ground truth) were classified by the gender-specific logistic classifier described above. Shown in Figure 5(a-b) are the resulting confusion matrices. The average accuracy of classifying a feminine body is 45.9%, as compared to 36.8% for masculine bodies. This classification accuracy stands in stark contrast to the 95% accuracy for the default neutral poses.

In an effort to make the classifiers more resilient to the 3D modeling errors, an augmented classifier, trained on both the neutral body shapes and half of the newly estimated body shapes, yields a bump in testing accuracy to 52.1% (feminine) and 46.1% (masculine). The corresponding confusion matrices are shown in Figure 5(c-d). By comparison, chance accuracy for a 12-way classifier is 8.3%.

Because of the discrete nature of this classifier, the class-based accuracy may not necessarily reflect the practical utility of this approach. For feminine body shapes, for example, body-shape 1 is never confused with much smaller body shapes, such as 6, 8, or 12. On the other hand, perceptually similar body shapes are likely to be confused, such as 4 being confused with 7 (57% of the time), or 6 being confused with 8 (56% of the time).

Although not a perfect comparison, we next evaluate the distinctiveness of this body-shape categorization with a similar 12-way height/weight classification. We begin by partitioning heights between 149 and 191 cm (4'11" and 6'3"), and weights between 41 and 113 kg (90 and 250 lbs) into three height categories (small [149-163 cm], medium [163-177 cm], large [177-191 cm]), and four weight categories (small [41-59 kg], medium [59-77 kg], large [77-95 kg], and x-large [95-113 kg]).

The distinctiveness of these twelve categories is evaluated as follows. A random ground-truth height and weight is drawn from the above ranges from which measurement noise is added to yield a measured height and weight. This measurement noise is drawn from a Gaussian distribution

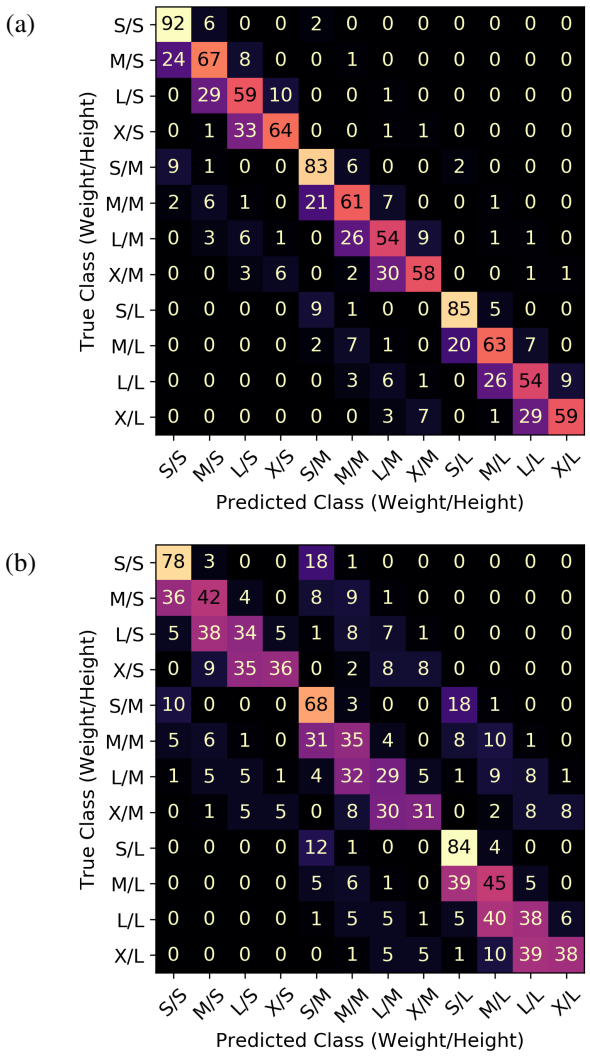

Figure 6. Confusion matrices evaluating weight/height distinctiveness using (a) CPD and (b) IPD scale disambiguation. Values are reported as percentages (values less than 0.5% are rounded down to 0). Class labels S,M,L refer to one of three height classes, and class labels S, M, L, X refer to one of four weight classes.

with mean and variance measured from our expected error in measuring height and weight, assuming a CPD or IPD scale disambiguation (see Section 4.3). The resulting measured height and weight is then classified based on the above 12-way categorization, and compared to ground-truth.

Shown in Figure 6 are the confusion matrices for CPD and IPD scale disambiguation. For the more accurate CPD, the average classification accuracy is 66.6%, as compared to 46.5% for IPD. By comparison, chance accuracy for a 12-way classifier is 8.3%. With perfect scale (CPD), this classification accuracy is significantly better than the body-shape classification. With imperfect scale, however, this accuracy is similar to the body-shape classification.

6. Discussion

Although our results suggest a comparable reliability for the scale-dependent height/weight and the scale-agnostic body-shape forensic identification, a truly direct comparison is nontrivial. First, as compared to the body-shape estimation, measurements of height/weight require an estimation of absolute scale, the accuracy of which can be highly variable. Second, due to the constraints of the SMPLX modeling, body-shape estimation requires knowledge of perceived gender. And third, the real-world distribution of body shapes is unknown, and therefore we don't know how the confusion of certain body shapes will translate into a real-world forensic setting. On balance, however, it appears that the these two techniques provide a similar – albeit imperfect – level of discrimination. It remains to be seen if a combination of these two techniques can provide more reliable identification.

We only evaluated these forensic identification techniques in simulation. Our expectation, however, is that reliability will only decrease in real-world scenarios. Additional complexities will include converting from 3D model volume to weight in kilograms, and to determining a person's ground-truth body shape. Until reliability can be significantly improved in simulation, it is reasonable to conclude that these techniques are not yet equipped to handle more complex real-world scenarios.

Many factors impact how these techniques will operate in practice. The human body, for example, is often obscured by clothing, making all measurements more difficult. It is also not clear if a different approach to building the body-shape equivalence classes could, in practice, lead to more distinctiveness. For example, a more fine-grained metric than IoU or a metric based on human perception of body shape, may yield more distinct body-shape categories. And, the presence of multiple images from different cameras or a video sequence could lead to improved accuracy.

In its current form, however, it would appear that accurate estimation of height, weight, and body shape remains a challenging problem, particularly for the purpose of an often high-stakes forensic identification.

7. Broader Implications

In 2006, the National Academy of Sciences (NAS) launched a large-scale analysis of the state of forensic science. Published three years later, the far-reaching 328-page report [9] called for a fundamental restructuring of how forensic techniques are validated and applied, and how forensic analysts are trained and accredited. One of the report's key findings was that "[w]ith the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source." The report argued that forensic practitioners too often offered evidence based on forensic techniques that had been shown to be invalid or unreliable, and that many forensic examiners exaggerated their testimony, inflating the reliability of their methods and conclusions.

A decade after the report's release, Judge Harry Edwards, co-chair of the original committee, wrote "we are still struggling with the inability of courts to assess the efficacy of forensic evidence. When a forensic expert testifies about a method that has not been found to be valid and reliable, the expert does not know what he does not know and cannot explain the limits of the evidence. This is unacceptable." [15]

As noted by the NAS, flawed forensic science can have severe consequences. the National Registry of Exonerations, for example, identified that, between 1989 and 2019, flawed or misleading evidence gathered using forensic techniques contributed to almost a quarter of wrongful convictions in the U.S. We contend, therefore, that significantly more attention should be paid to understanding the limits of, and improving the reliability of, forensic identification techniques.

References

[1] NHANES questionnaires, datasets, and related documentation. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2015. Accessed: 2022-2-22. 4

[2] Ivo Alberink and Annabel Bolck. Obtaining confidence intervals and likelihood ratios for body height estimations in images. *Forensic Science International*, 177(2-3):228–237, 2008. 2

[3] Olasimbo Ayodeji Arigbabu, Sharifah Mumtazah Syed Ahmad, Wan Azizun Wan Adnan, Salman Yussof, Vahab Iranmanesh, and Fahad Layth Malallah. Estimating body related soft biometric traits in video frames. *The Scientific World Journal*, 2014, 2014. 2

[4] C. BenAbdelkader and Y. Yacoob. Statistical body height estimation from a single image. In *IEEE International Conference on Automatic Face Gesture Recognition*, pages 1–7, 2008. 2

[5] Surabhi Bhutani, Eva Kahn, Esra Tasali, and Dale A Schoeller. Composition of two-week change in body weight under unrestricted free-living conditions. *Physiological Reports*, 5(13):e13336, 2017. 1

[6] Didier Bieler, Semih Günel, Pascal Fua, and Helge Rhodin. Gravity as a reference for estimating a person's height from video. *International Conference on Computer Vision*, pages 8568–8576, 2019. 2

[7] Zhe Cao, T. Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1302–1310, 2017. 3

[8] Rama Chellappa and Pavan Turaga. Recent advances in age and height estimation from still images and video. In *Face and Gesture*, pages 91–96, 2011. 2

[9] National Research Council Committee on Identifying the Needs of the Forensic Sciences Community. *Strengthening forensic science in the United States: A path forward*. National Academies Press, 2009. 8

[10] A. Criminisi, I. Reid, and Andrew Zisserman. Single view metrology. *International Journal of Computer Vision*, 40:123–148, 2004. 2

[11] Antonio Criminisi, Andrew Zisserman, Luc J. Van Gool, Simon K. Bramble, and David Compton. New approach to obtain height measurements from video. In Kathleen Higgins, editor, *Investigation and Forensic Science Technologies*, volume 3576, pages 227 – 238. International Society for Optics and Photonics, 1999. 1, 2

[12] Neil A Dodgson. Variation and extrema of human interpupillary distance. In *Stereoscopic Displays and Virtual Reality Systems XI*, volume 5291, pages 36–46. International Society for Optics and Photonics, 2004. 4

[13] Yujie Dong and Damon L Woodard. Eyebrow shape-based features for biometric recognition and gender classification: A feasibility study. In *International Joint Conference on Biometrics*, pages 1–8, 2011. 3

[14] Nicolae Duta. A survey of biometric technology based on hand shape. *Pattern Recognition*, 42(11):2797–2806, 2009. 3

[15] Harry T Edwards. Ten years after the National Academy of Sciences' landmark report on strengthening forensic science in the United States: A path forward–where are we? *Available at SSRN 3379373*, 2019. 8

[16] Dale A Gerke, Jean-Michel Brismée, Phillip S Sizer, Gregory S Dedrick, and C Roger James. Change in spine height measurements following sustained mid-range and end-range flexion of the lumbar spine. *Applied Ergonomics*, 42(2):331–336, 2011. 1

[17] Afzal Godil, Patrick Grother, and Sandy Ressler. Human identification from body shape. In *International Conference on 3-D Digital Imaging and Modeling*, pages 386–392, 2003. 3

[18] Enrique Gómez, Carlos M Travieso, Juan C Briceño, and Miguel A Ferrer. Biometric identification system by lip shape. In *International Carnahan Conference on Security Technology*, pages 39–42, 2002. 3

[19] Ester Gonzalez-Sosa, Ruben Vera-Rodriguez, Julian Fierrez, and Javier Ortega-Garcia. Comparison of body shape descriptors for biometric recognition using MMW images. In *International Conference on Pattern Recognition*, pages 124–129, 2014. 3

[20] Mohamed Hassan, Vasileios Choutas, Dimitrios Tzionas, and Michael J. Black. Resolving 3D human pose ambiguities with 3D scene constraints. In *International Conference on Computer Vision*, oct 2019. 1

[21] Anil K Jain and Stan Z Li. *Handbook of face recognition*, volume 1. Springer, 2011. 1

[22] Avinash Kalyanaraman, Dezhi Hong, Elahe Soltanaghaei, and Kamin Whitehouse. Forma track: tracking people based on body shape. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):1–21, 2017. 3

[23] Hiroharu Kato, Yoshitaka Ushiku, and Tatsuya Harada. Neural 3D mesh renderer. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2018. 3

[24] Christoph Lassner, Javier Romero, Martin Kiefel, Federica Bogo, Michael J Black, and Peter V Gehler. Unite the people: Closing the loop between 3D and 2D human representations. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 6050–6059, 2017. 1

[25] Dong-Seok Lee, Jong-Soo Kim, Seok Chan Jeong, and Soon-Kak Kwon. Human height estimation by color deep learning and depth 3D conversion. *Applied Sciences*, 10(16):5531, 2020. 2

[26] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. SMPL: A skinned multi-person linear model. *ACM Transactions on Graphics (SIGGRAPH Asia)*, 34(6):248:1–248:16, 2015. 1, 3

[27] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32:2262–2275, 2010. 4

[28] Tam V Nguyen, Jiashi Feng, and Shuicheng Yan. Seeing human weight from a single RGB-D image. *Journal of Computer Science and Technology*, 29(5):777–784, 2014. 2

[29] Angela Olver, Helen Guryn, and Eugene Liscio. The effects of camera resolution and distance on suspect height analysis using PhotoModeler. *Forensic Science International*, page 110601, 2020. 2

[30] Georgios Pavlakos, Vasileios Choutas, Nima Ghorbani, Timo Bolkart, Ahmed AA Osman, Dimitrios Tzionas, and Michael J Black. Expressive body capture: 3D hands, face, and body from a single image. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 10975–10985, 2019. 1, 2, 3, 4

[31] Neerja Thakkar and Hany Farid. On the feasibility of 3D model-based forensic height and weight estimation. In *IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Media Forensics*, pages 953–961, 2021. 1, 2, 3, 4

[32] Carmelo Velardo and Jean-Luc Dugelay. Weight estimation from visual body appearance. In *IEEE International Conference on Biometrics: Theory, Applications and Systems*, pages 1–6, 2010. 2

[33] Ping Yan and Kevin W Bowyer. Biometric recognition using 3D ear shape. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(8):1297–1308, 2007. 3

[34] Rui Zhu, X. Yang, Yannick Hold-Geoffroy, Federico Perazzi, J. Eisenmann, Kalyan Sunkavalli, and M. Chandraker. Single view metrology in the wild. *arXiv*, abs/2007.09529, 2020. 2