

Jeremy N. Bailenson
bailenson@stanford.edu
Virtual Human Interaction Lab
Department of Communication
Stanford University
Stanford, CA 94305-2050

Andrew C. Beall
beall@psych.ucsb.edu

Jim Blascovich
blascovich@psych.ucsb.edu

Christopher Rex
crex@uic.edu
Research Center for Virtual
Environments and Behavior
Department of Psychology
University of California at Santa
Barbara
Santa Barbara, CA 93106

Examining Virtual Busts: Are Photogrammetrically Generated Head Models Effective for Person Identification?

Abstract

We examined the effectiveness of using 3D, visual, digital representations of human heads and faces (i.e., virtual busts) for person identification. In a series of 11 studies, participants learned a number of human faces from analog photographs. We then crafted virtual busts from those analog photographs, and compared recognition of photographs of the virtual busts to the original analog photographs. We demonstrated that the accuracy of person identification using photographs of virtual busts is high in an absolute sense, but not as high as using the original analog photographs. We present a paradigm for comparing the similarity, both structural (objectively similar in shape) and subjective (subjectively in the eyes of a viewer) of virtual busts to analog photographs, with the goal of beginning the discussion of a uniform standard for assessing the fidelity of digital models of human faces.

I Introduction

In ancient Greece, aristocrats paid sculptors large sums to craft realistic marble busts of themselves, and only the wealthiest could afford them. Today, however, digital technology makes possible inexpensive crafting of accurate digital or virtual 3D representations of heads and faces (i.e., “virtual busts”) from a series of analog photographs. Such digital technology provides scholars and researchers tools for uses such as animation, historical preservation, and, most relevant to the purposes of this paper, “person identification.”

Person identification involves psychological processes underlying the recognition and identification of specific human faces and bodies. Person identification has been studied by many behavioral science researchers (Cutting & Kozlowski, 1977; Malpass & Devine, 1981; Krafka & Penrod, 1985; Bruce, Valentine, & Baddeley, 1987; Brigham, Meissner, & Wasserman, 1999; Wells, 2002), forensic artists (Prag, Neave, & Neave, 1997; Taylor, 2000), medical doctors (Farkas, 1981), and computer scientists (Brunelli, Falavigna, Poggio, & Stringa, 1995; BenAbdelkader, Cutler, & Davis, 2002). Most of this work demonstrates that individuals frequently make errors during person identification, especially during eyewitness testimony (see Wells for a review).

Person-identification-based research can potentially benefit from the use of digital busts given their distinct advantages over normal (i.e., analog) photographs (Bailenson, Beall, & Blascovich, 2003; Bailenson, Wiggins, Blascovich,

& Beall, 2003). First, 3D digital representations permit identifiers to examine the represented person from any arbitrary angle and distance, including the angle and distance from which he or she may have originally seen the person. Second, it is trivial to render the representation stereoscopically, thus providing binocular cues (Liu, Collin, & Chaudhuri, 2000) to depth that provide more information than photographs and greatly enhance realism. Third, using animation sequences, it is possible to digitally render dynamic (i.e., animated) facial gestures and spoken actions that may be crucial for identification (Bruce & Valentine, 1988). Consequently, forensic artists, investigators, legal counselors, and intelligence agencies all can potentially derive value from digital busts. The purpose here was to examine the utility of 3D models of heads and faces for person identification.

We created models by taking photographs of people and then using commercially available photogrammetric software (3DMeNow, version 1.5) to build 3D models digitally. Photogrammetric software (e.g., 3DMeNow, FaceStation, FaceGen) allows creation of digital virtual busts from two photographs: a full (i.e., front-on) view and a profile view. The developer scans the photographs, and the software then computes the underlying 3D head structure by means of the two unique views provided from the photographs. The software permits the user to adjust the fitted model manually to better perfect the match between the digital model and the individual's true head shape. Typically, the manual adjustments are relatively small.

This technique is an alternative to a 3D scanner, which uses optical methods to directly measure the shape and albedo of a person's head and face. Software such as 3DMeNow has two distinct advantages over the scanner: it is relatively inexpensive compared to most 3D scanners, and the represented person does not need to be physically present when the software is being used (only two photographs are necessary). However, these benefits may come at a functional cost, since the 3D scanner should produce a more accurate representation than the photogrammetric software because the scanner directly measures points on a person's head and face at a

very high resolution. This contrasts with typical photogrammetric methods, which parametrically align a model to the image data using relatively few control points.

In the current set of studies, we examined how accurately experimental participants could recognize identities from virtual photographs taken of our photogrammetrically reconstructed virtual busts compared to the original photographs. Given that there is a loss, albeit minimal, of structural accuracy that results from deriving a 3D shape from two photographs in this manner (i.e., the model is not built *exactly* like the represented person's head and face), we sought to determine the extent of any inaccuracies in terms of person identification performance. Furthermore, we examined sources of the structural inaccuracies of the 3D models by employing objective measures of similarity assessment. In the following sections, we describe a series of 11 studies that test the utility of these virtual models for recognition and identification. We also characterize the underlying differences between photogrammetric digital models and the photographs on which they were based in terms of subjective and structural similarity.

2 Experiments

2.1 Recognition Experiments

In the set of 11 studies, participants were exposed to photographs of human faces and subsequently tested on their recognition of those faces from different angles (i.e., full, three-quarter view, or profile) as well as from different media (i.e., analog photographs or virtual photographs of digital models) presented via computer monitor. The goal of these studies was to examine participants' recognition of digital faces and to compare this performance to recognition of analog photographed faces (Troje & Bueltoff, 1996) from different angles and contexts.

2.1.1 Common Methods. We utilized a face database consisting of photographs of 22 Caucasian

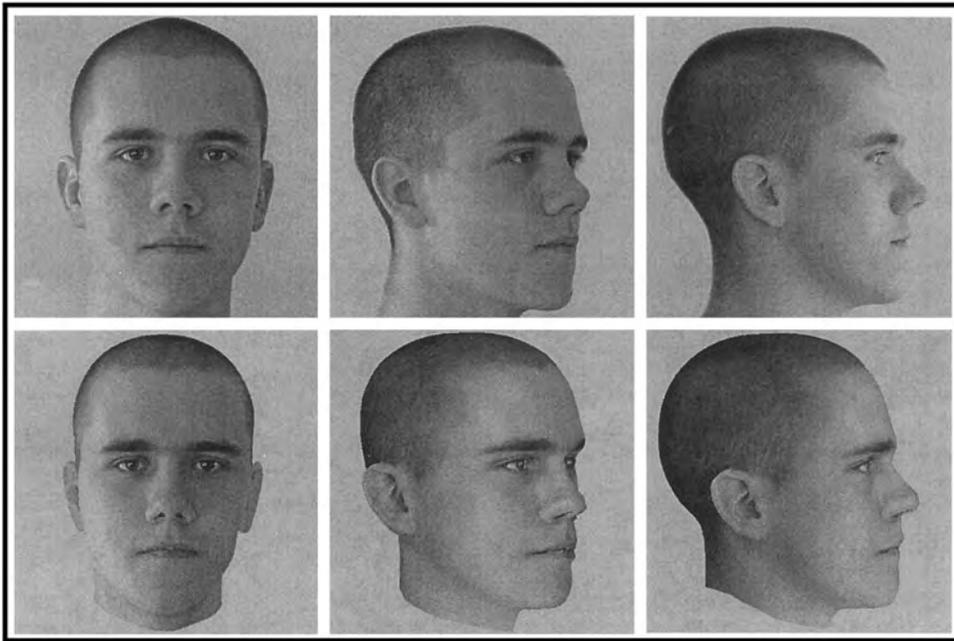


Figure 1. A sample stimulus from the recognition studies. The top row contains photographs; the bottom row contains screen-shots of the virtual model. Full views are on the left, three-quarter views are in the middle, and profile views are on the right.

males from a number of different angles as stimuli.¹ Based on frontal and profile images, we used modeling software (3dMeNow) to create a 3D, digital model (i.e., virtual busts) from each image pair. Consequently, we had 22 digital virtual busts. The manual adjustments necessary to create each bust took approximately forty minutes. For each virtual bust, we then created virtual photographs (i.e., “snapshots”) matching the same angles as the original analog photographs. Figure 1 shows the heads at different angles for both photographs and digital busts for a single face.

There were 191 participants across all studies who received partial course credit for an introductory psychology course. Participants were tested alone.

In the exposure, or training session, we presented each participant analog photographic images of 11 faces

randomly selected from our database. We randomly paired each face with an arbitrary first name in order to increase exposure time to the images and to motivate participants to examine the images more closely. The names were only used during this initial training session. Participants scrolled through each face, one at a time, until they had learned the names of the set of 11 faces to an 80% correct criterion.

After the training session, participants began the recognition testing session. We manipulated and crossed two within-subject variables, Image type (analog photograph or digital model) and Face Angle (full, three-quarter, or profile) resulting in six different conditions; each participant experienced all six conditions.

The recognition test was forced choice.² During each

1. The faces were utilized courtesy of the Computer Vision Laboratory, Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia and SCV, PTERS, Velenje, Slovenia, <http://www.lrv.fri.uni-lj.si/facedb.html>

2. We chose a forced-choice paradigm for statistical reasons and point out that other applications, such as witness identification, may be better served by a yes-no paradigm, which allows the observers an added degree of freedom to set his or her decision criterion level (e.g., confidence).

Table 1. Summary of Experimental Contexts for the 11 Recognition Studies

Study	Trained on	Training time	Testing time	Feedback	Sample size	Hair	RV difference (%)
1	Full real	Unlimited	Unlimited	Yes	20	Yes	9
2	Full real	Unlimited	300 ms	Yes	14	Yes	12
3	Full real	1000 ms	Unlimited	Yes	17	Yes	7
4	Full real	Unlimited	Unlimited	No	15	Yes	12
5	Full real	Unlimited	Unlimited	Yes	21	No	3
6	Full real	Unlimited	Unlimited	Yes	17	During training	5
7	Full virtual	Unlimited	Unlimited	Yes	18	Yes	4
8	Full virtual	Unlimited	Unlimited	No	19	Yes	2
9	Profile virtual	Unlimited	Unlimited	Yes	12	Yes	-5
10	Profile real	Unlimited	Unlimited	Yes	21	Yes	15
11	Three-quarter real	Unlimited	Unlimited	Yes	17	Yes	17

Note. In every condition across all studies, recognition was significantly greater than chance (over 50%).

trial, participants saw two faces. Both images in each pair were of the same type (e.g., digital virtual profiles were paired with one another), and all pairs contained one face that participants had learned during the training session and one unfamiliar distracter. Pairs were presented side by side on one computer monitor. Participants were instructed to indicate which of the two faces was one they had been exposed to during the training session. Selection was made by pressing a key on the computer keyboard. The faces remained on the screen until participants made their choice, after which a new pair of faces appeared.

For each training face, participants were tested in all six conditions. Consequently there were 66 unique trials in each block. Participants were tested on a total of four blocks (264 trials total). The order of trials within each block was random, and each block featured exactly the same trials, with the same test faces paired with the same distracter faces. There was no break between blocks. Furthermore, appearance of test faces and distracter faces on screen-side (i.e., left vs. right) was random.

2.1.2 Methods of Specific Studies. The purpose of running 11 different studies was to gauge the difference between analog photographs and digital virtual busts across a variety of psychological contexts, including exposure time to the stimuli (this variable was

chosen to examine how automatic facial features were encoded; small exposures simulate single glances while longer exposures simulate more extensive exposure), the presence of feedback (this variable was chosen to examine the relative influences of pure memory in the no-feedback condition and the combination of memory and learning in the feedback condition), and the amount of information portrayed in the images (this variable was chosen in order to examine the extent to which including hair, which is notoriously difficult to model, hampers recognition). These manipulations have been established previously as important parameters in studying recognition (Tulving, 1983). Furthermore, these different contexts may play an important role in situations related to identification (Wells, 2002). We implemented extensive pilot tests to determine the specific levels of these variables, such as the duration of the reaction time trials and the types of feedback, and chose levels that kept performance away from floor and ceiling levels. Across studies, we did not exhaust every combination of these variables; instead, based on the results of a completed study, we manipulated variables in the next study that seemed psychologically important.

Different groups of people participated in each study. Table 1 describes these different contexts by providing the details concerning methodology for the 11 studies. The second column of Table 1, “Trained on,” indicates

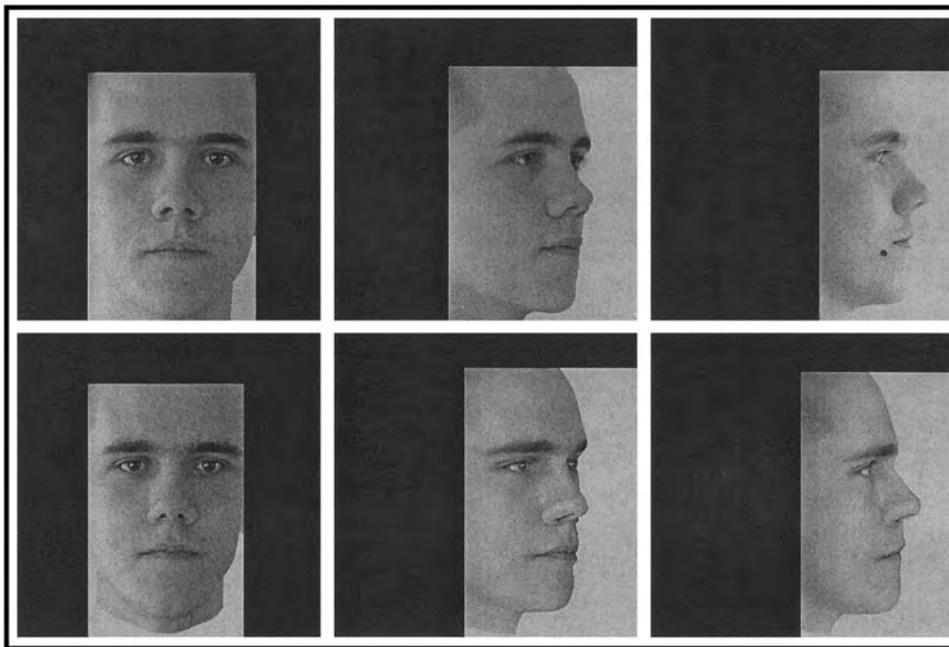


Figure 2. A sample stimulus of images with the hair blocked out. The top row contains photographs; the bottom row contains screen-shots of the virtual model.

the angle and media of the faces on which participants were trained. The third column, “Training time,” indicates the amount of time for which participants were exposed to each training face during the training session. “Testing time” denotes the amount of time for which participants were exposed to each pair of test images before making their recognition decision. “Feedback” indicates whether or not participants received appropriate feedback for correct and incorrect responses during the testing session. “Sample size” denotes the number of participants from each study. “Hair” indicates whether participants viewed complete images of the stimuli or partial images with the hair blocked out, as Figure 2 illustrates.

2.1.3 Results and Discussion. The general results appear in the rightmost column of Table 1. “RV difference” indicates the difference in percentage in recognition (chance is 50%, perfect score is 100%) between the analog photographs and the digital images. Positive scores indicate that participants could more easily recog-

nize training faces from real photographs than from images of virtual busts.

In all studies, recognition of faces from all angles and media were reliably above chance (50%) according to paired t -tests with an alpha of .01. Across the 11 studies, the mean difference between analog photographic and digital virtual recognition was approximately seven percent. This difference was statistically different from zero, $t(190) = 11.354$, $p < .001$,³ such that participants were generally more adept at recognizing training faces from analog photographs than from snapshots of virtual models.

A previous face recognition finding indicates that faces are more difficult to recognize in direct proportion

3. In certain statistical analyses in this section, we collapsed across studies and pooled participants from different experiments together. This is somewhat problematic in terms of random assignment, in that participants across studies were run at different times. However, the statistically reliable effects also occur in a by-experiment basis for all studies that include hair; grouping across studies provides clarity for the reader. Furthermore, these individual effect sizes are quite large.

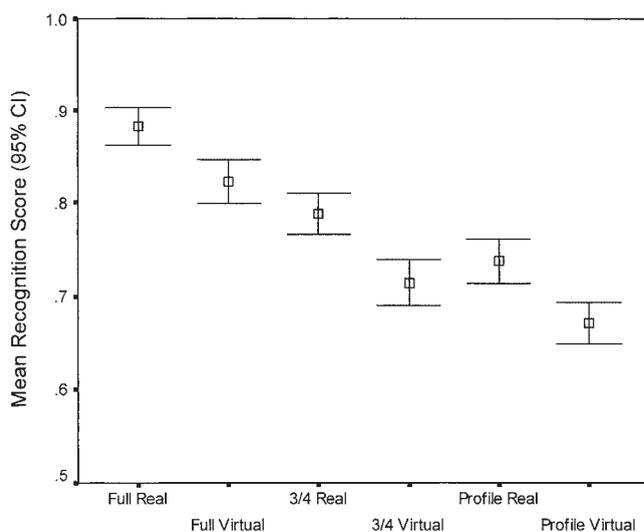


Figure 3. Recognition performance from the six experimental conditions from studies on which participants were trained on the full view of the faces.

to the difference between training angle and testing angle (Troje & Buelhoff, 1996). In other words, when participants learn a face at a given angle, it is difficult to recognize that face when it is presented at angles far from the original angle. Of interest in the current study is whether or not this trend occurs for our virtual busts as well as for photographs of actual heads. Figure 3 shows the mean recognition performance for all participants in the first eight studies described in Table 1, the ones in which participants were trained with the frontal or “full” view (i.e., the angle during training was 0). The data demonstrate a clear and significant linear trend for both analog photographic faces ($F(1,140)=157.45$, $p < .001$) and for digital faces ($F(1,140)=138.13$, $p < .001$). In other words, for virtual busts, as well as for analog photographs, it is more difficult to recognize faces when the angle during testing is far from the angle during training than when the two angles are similar.

Hair is recognized as a notoriously difficult feature to model digitally. As Table 1 illustrates, when we removed the hair features from the face images, the difference between analog photographic and virtual face recognition greatly diminishes. The average difference between photographic and digital in studies 5 and 6

(where hair features were absent) was only 4% compared to 9% for the other studies. This difference was significant, $F(1, 189) = 7.01$, $p < 0.01$. These results suggest that in general, participants may have demonstrated poorer recognition for virtual faces than for real faces in part due to the difficulty of veridically digitally modeling hair.

In these studies, we demonstrated four important findings. First, participants were able to discriminate learned, digitally modeled faces from distracter, digitally modeled faces reliably above chance. Second, the well-established degradation that occurs as the difference in head orientation between learning and testing increases occurred for our digitally modeled faces, indicating that, on some levels, people process these faces in a similar manner to analog photographs of faces. Third, still images of our digital busts are not as effective as analog photographs in terms of person identification in which participants are trained on 2D photographs. Finally, these three findings persist across a variety of experimental contexts and memory conditions.

Still images of digital busts did not aid in person identification as much as analog photographs. This is not surprising considering that: (a) the digital busts were crafted from the exact photographs that were later tested against the busts, (b) participants were trained on 2D images and consequently could not take advantage of depth cues and integration from multiple angles, and (c) there is some functional cost in accuracy with the photogrammetric process. It is quite notable, given these three factors, that the difference between the virtual busts and the analog photographs was consistently small. In the following sections, we use structural analyses and subjective ratings as a way to more systematically determine the sources of these dissimilarities between the real and virtual heads responsible for the differences in recognition.

2.2 Real/Virtual Similarities

We utilize two distinct tools to gauge the similarity between the digital virtual images and the analog photographs on which those models were based. The first is *anthropometric similarity*, a structural measure

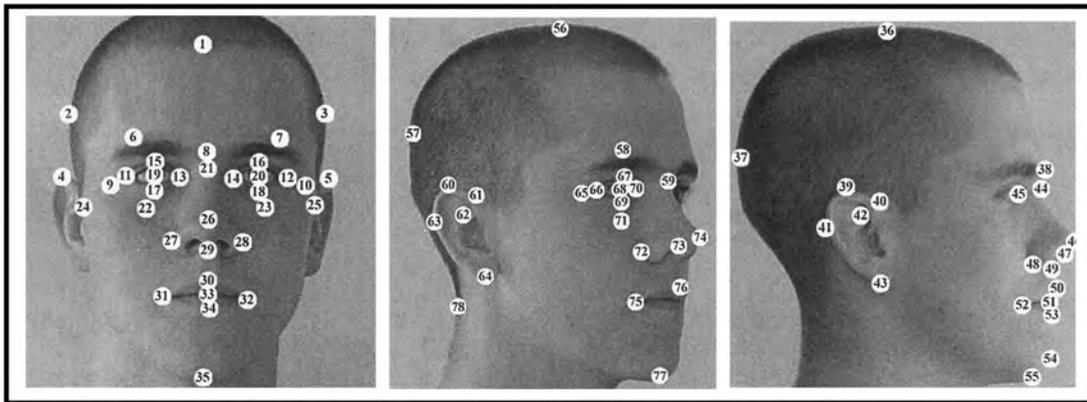


Figure 4. The anthropometric landmarks described in Appendix I. There are some redundant landmarks across view angles. Those points have unique numbers. In this figure, the size of the points is exaggerated to give the reader a rough idea of their location. See Farkas (1981) for a more fine-tuned description of these landmarks.

based on the positioning of certain known landmark points on the head and face. The second is *subjective similarity*, a measure taken from a separate group of participants who viewed the analog photographic and digital virtual faces and reported their similarity on a ratings scale.

2.2.1 Anthropometrics. Anthropometric analyses have been utilized by psychologists to study face perception (Busey, 1988) and by computer scientists to develop digital virtual models of the head and face (Decarlo, Metaxas, & Stone, 1998; Blanz & Vetter, 1999). Furthermore, based on anthropometrics, researchers are developing automatic detection of facial landmarks using neural networks (Bartlett et al., 2000). In the current analysis, we used a set of landmark points roughly based on a medical measurement procedure developed by Farkas (1981); Figure 4 shows the landmark points used on each head view, and the appendix provides the name and a short description of each landmark point.

In order to produce a measure of anthropometric similarity, we located each of those landmark points on all 22 unique faces utilized in the reported studies for each of the six unique versions of each face (i.e., full real, full virtual, three-quarter real, three-quarter virtual, profile real, profile virtual). First, we calculated all the pairwise differences for each of the landmarks shown in

Figure 4 between analog photographic and digital faces. Next, we normalized these differences using the largest pairwise difference (between points 1 and 35 for full, 56 and 77 for three-quarter, and 36 and 55 for profile) as the denominator for each difference. Then, for each angle, we computed the difference of all corresponding pairwise distances for the analog photographic and digital virtual version. Finally, we summed the absolute value of all those differences. Consequently, for each angle on each unique face, we computed a single measure (i.e., anthropometric mismatch) that represented the total difference of landmark distances between the photographic and digital heads.

Overall, the digital heads were quite similar to the photographic heads. The correlation in an item analysis between photographic distances and digital head distances (i.e., all the pairwise distances for all of the angles on all unique faces: 22,836 observations) was almost perfect ($r = .987, p < 0.001$). However, this high correlation is due in part to the natural similarities of human facial structure in general across target heads (correlations of the real distances from one analog head and the same virtual distances of a different digital head are generally over .95). Figure 5 shows the mean anthropometric mismatch (i.e., the difference between real distances and virtual distances), comparing pairwise distances that are unique to each angle.

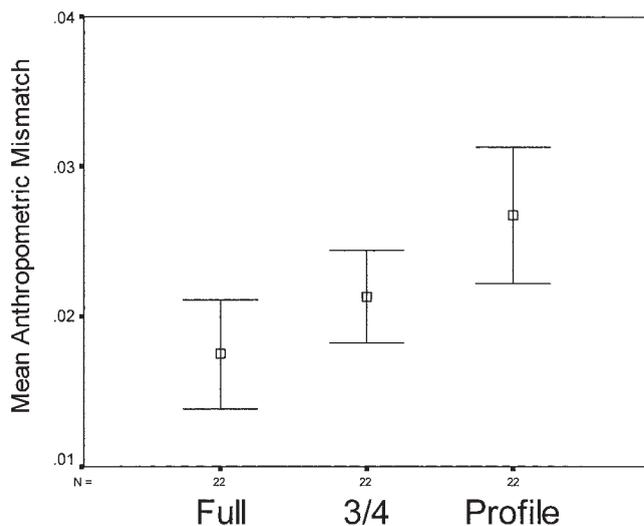


Figure 5. Mean anthropometric mismatch for the 22 heads by angle.

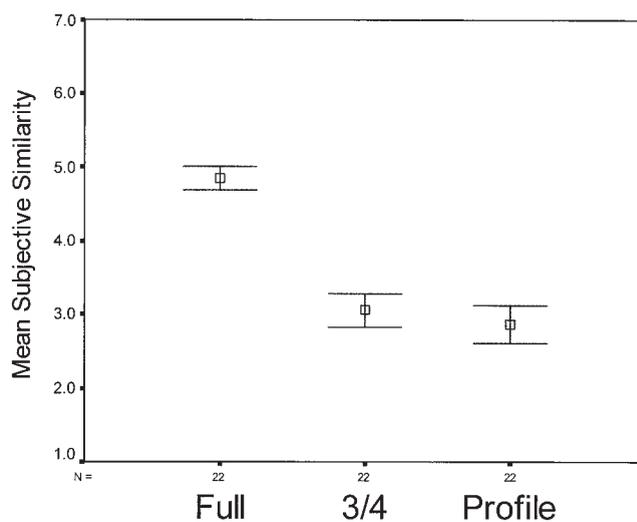


Figure 6. Mean subjective similarity for the 22 heads by angle.

We had originally predicted no differences in anthropometric mismatch across the three angles (at least between full and profile) because there was no reason to suspect that the software would treat those angles differently. However, we conducted an item-analysis ANOVA with the number of unique possible trial types as the error term (66 observations: 3 views of 22 heads), angle as the independent variable, and anthropometric mismatch as the dependent variable. The effect of angle was found to be significant, $F(2, 63) = 6.51, p < 0.005$. Post-hoc tests using Fisher's LSD indicate that all pairwise differences are reliable with an alpha of .05 except for the difference between full and three-quarter. In sum, the photogrammetric models are least structurally accurate (i.e., have the most anthropometric mismatch) at the profile. This finding is odd, considering that the profile should be more accurate than the three-quarter view, since the profile is largely based on an original photograph. We believe that this effect may have to do with most profiles containing large amounts of hair (compared to the other views).

2.2.2 Subjective Similarity. The purpose of measuring subjective similarity was to acquire some measure of phenomenological similarity between the analog photographic and digital heads in order to deter-

mine how sensitive people are to structural similarity when examining digital virtual heads. To measure subjective similarity we ran an experiment in which 16 participants rated the similarity of the two image types (digital vs. analog) for all 66 pairs of faces (3 angles of 22 faces). Each participant rated each pair twice, once with the digital head on the left and once with it on the right. The pairs were displayed on a computer monitor in a random order, and participants rated similarity on a seven point scale with seven indicating "extremely similar" and one indicating "not similar at all." Figure 6 shows the mean subjective similarity by viewing angle.

We ran another ANOVA with "item" as the error term, "angle" as the independent variable, and subjective similarity as the dependent variable. The effect of angle was significant, $F(2, 63) = 108.89, p < 0.0001$. Post-hoc tests using Fisher's LSD indicate the full angle is different from the other two angles, but three-quarter is not reliably different from profile. In sum, participants perceived the photogrammetric digital models to be most similar to the analog photographs when viewed at the full angle. Furthermore, the correlation of subjective similarity with anthropometric mismatch (66 observations) was reliable, $r = 0.29, p < 0.05$, such that the higher the mismatch, the less similar the faces were rated. However, this correlation is low, and subjective

similarity is by no means dictated by structural similarities. Instead, participants may be attending to some set of second-order combinations of these structural features.

2.2.3 Similarity and Recognition. The similarity data provide a potential explanation for the recognition data. In other words, we can test to see how large a role structural and subjective similarity play in transfer of recognition from analog photographs to digital virtual models. To do so, we looked at the recognition performance of each unique face (i.e., 22 unique faces at three angles). For each one, we computed the average RV difference (i.e., the last column in Table 1) across the eight studies described above on which participants were trained on photographic faces and tested on digital virtual faces. Consequently, for each unique face we had an average measure that indicated how well recognition transferred from analog photographic to digital virtual models. We then ran an ANCOVA with mean RV difference as the independent variable, view angle as a fixed factor, and anthropometric and subjective similarity as covariates. The only significant effect was subjective similarity, $F(1, 61) = 7.41, p < 0.01$, such that the more similarly analog photographic and virtual heads were rated, the smaller the RV difference in recognition.

This analysis demonstrates that structural similarity is less valuable in predicting the recognition of virtual busts than subjective similarity ratings. In other words, our participants making subjective ratings are attending to features of the heads that are not captured in our anthropometric analysis. In future work, we plan on exploring what these features may be.⁴

3 Conclusions

We examined people's recognition of virtual busts produced with photogrammetric software. While our

4. We implemented a factor analysis to determine which structural features clustered together, and then used those clusters to predict subjective similarity in a regression. However, the results of this analysis did not reveal any significant or noteworthy trends.

data show that people are able to reliably recognize photographs of the digital virtual busts of people after being exposed to analog photographs of the same people, they do so with a small cost in accuracy (less than 10%) compared to recognizing the original photographs. Furthermore, subjective similarity ratings were more effective at predicting the loss in accuracy than an objective anthropometric measure of similarity.

These results are encouraging for the prospect of using digital virtual busts to aid person identification. The virtual busts used in the current set of studies were produced at minimal cost in terms of time (about 1 hr per head) and money (off-the-shelf software). Better allocation of resources in crafting a given digital virtual bust should result in better recognition. Furthermore, in the coming years, technology for producing digital virtual busts will likely improve, and the differences in recognition found here should diminish. Digital virtual busts should prove to be an extremely helpful tool to aid in the study and implementation of person identification.

We also presented a paradigm for assessing the utility of virtual busts for person identification involving recognition performance, anthropometrics, and subjective similarity. The current research is limited in that we only examine photogrammetrically-generated virtual busts. In future research, we plan on using this paradigm to examine other methods of producing virtual busts, such as with a 3D scanner.

However, the methodological paradigm raised in the current paper should function as a substantial beginning in the process of developing a uniform standard for the assessment of virtual bust effectiveness. Digital human representation is becoming more commonplace in a variety of contexts, including movies, communication systems, chatrooms, and online video games. A standardized array of methods for identifying a virtual bust will be not only helpful, but essential, once the uses of virtual busts become more prevalent in situations where the stakes are high (e.g., person identification, monetary transactions, etc.) In the current paper, we demonstrate a starting point for such a uniform standard. However, much work needs to be done towards this goal, including examining virtual busts made in different fashions as well as working towards an absolute, objective standard.

Acknowledgments

The authors would like to thank Russell Terry for assistance in data analysis as well as Lauren Dahl, Natalie Fabert, Leila Gorimar, Marin Howell, Robert Mathrole, Jason Mesick, and Natalie Teague for assistance in collecting data. This research was sponsored in part by NSF Award SBE-9873432 and in part by NSF ITR Award 0219399.

References

- Bailenson, J. N., Beall, A. C., & Blascovich, J. (2003). Using virtual heads for person identification: An empirical study comparing photographs to photogrammetrically-generated models. *Journal of Forensic Identification*, 53(6), 722–728.
- Bailenson, J. N., Blascovich, J., & Beall, A. C. (2004). Courtroom applications of virtual environments, immersive virtual environments, and shared virtual environments. *Law and Policy*. Available on the web at <http://www.blackwellpublishing.com/journal.asp?ref=0265-8240&site=1>
- Bartlett, M. S., Donato, G. L., Movellan, J. R., Hager, J. C., Ekman, P., & Sejnowski, T. J. (2000). Image representations for facial expression coding. In S. Solla, T. Leen, & K. Mueller (Eds.), *Advances in neural information processing systems*, 12 (pp. 886–892). Cambridge, MA: MIT Press.
- BenAbdelkader, C., Cutler, R., & Davis, L. (2002). Person identification using automatic height and stride estimation. *Proceedings of IEEE International Conference on Pattern Recognition 2002*.
- Blanz, V., & Vetter T. (1999). A morphable model for the synthesis of 3D faces. *Proceedings of SIGGRAPH'99*, 187–194.
- Brigham, J. C., Meissner, C. A., & Wasserman, A. W. (1999). Applied issues in the construction and expert assessment of photo lineups. *Applied Cognitive Psychology*, 13, S73–S92.
- Bruce, V., & Valentine, T. (1988). When a nod's as good as a wink: The role of dynamic information in face recognition. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (vol. 1; pp. 169–174). Chichester, UK: John Wiley & Sons.
- Bruce, V., Valentine, T., & Baddeley, A. D. (1987). The basis for the 3/4 view advantage in face recognition. *Applied Cognitive Psychology*, 1, 109–120.
- Brunelli, R., Falavigna, D., Poggio, T., & Stringa, L. (1995). Automatic person recognition by using acoustic and geometric features. *Machine Vision and Applications*, 8, 317–325.
- Busey, T. A. (1988). Physical and psychological representations of faces: Evidence from morphing. *Psychological Science*, 9, 476–483.
- Cutting, J., & Kozlowski, L. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9(5), 353–356.
- Decarlo, D., Metaxas, D., & Stone, M. (1998). An anthropometric face model using variational techniques. *Proceedings of SIGGRAPH'98*, 67–74.
- Farkas, L. G. (1981). *Anthropometry of the head and face in medicine*. North Holland: Elsevier.
- Krafka, C., & Penrod, S. (1985). Reinstatement of context in a field experiment on eyewitness identification. *Journal of Personality & Social Psychology*, 49(1), 58–69.
- Liu, C. H., Collin, C. A., & Chaudhuri, A. (2000). Does face recognition rely on encoding of 3-D surface? Examining the role of shape-from-shading and shape-from-stereo. *Perception*, 29, 729–743.
- Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. *Journal of Applied Psychology*, 66, 482–489.
- Prag, J., Neave, R., & Neave, R. (1997). *Making faces: Using forensic and archaeological evidence*. Texas A & M University Press.
- Taylor, T. T. (2000). *Forensic art and illustration*. Boca Raton, FL: CRC Press.
- Troje, N. F., & Buelthoff, H. H. (1996). Face recognition under varying pose: The role of texture and shape. *Vision Research*, 36, 1761–1771.
- Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.
- Wells, G. L. (2002). Eyewitness identification: Psychological aspects. *The Encyclopedia of Crime and Justice*. New York: Macmillan.

Appendix: Summary of Points

Unless noted with an asterisk, the landmark points were taken from Farkas (1981).

Full View Points

1. Trichion—point on the hairline in the middle of the forehead

2. Eurion (right)—most prominent lateral point on the temporal bone
 3. Eurion (left)—most prominent lateral point on the temporal bone
 4. Otobasion superius (right)—upper border of the ear region
 5. Otobasion superius (left)—upper border of the ear region
 6. Superciliare (right)—highest point of each eyebrow
 7. Superciliare (left)—highest point of each eyebrow
 8. Glabella—midline point between the eyebrows
 9. Frontosupraorbitale (right)—point where brow ridge meets ridge under the eye
 10. Frontosupraorbitale (left)—point where brow ridge meets ridge under the eye
 11. Exocanthion (right)—lateral hinge where the eyelid closes
 12. Exocanthion (left)—lateral hinge where the eyelid closes
 13. Endocanthion (right)—medial hinge where the eyelid closes
 14. Endocanthion (left)—medial hinge where the eyelid closes
 15. Palpebrale superius (right)—highest point of the upper eyelid line
 16. Palpebrale superius (left)—highest point of the upper eyelid line
 17. Palpebrale inferius (right)—lowest point of the lower eyelid line
 18. Palpebrale inferius (left)—lowest point of the lower eyelid line
 19. Center point of the pupil (right)*
 20. Center point of the pupil (left)*
 21. Nasion—most indented point on the nose ridge within the orbital region
 22. Orbitalis (right)—bony ridge below the eye
 23. Orbitalis (left)—bony ridge below the eye
 24. Zygion (right)—most lateral point of the zygomatic arch; at the level of the ears
 25. Zygion (left)—most lateral point of the zygomatic arch; at the level of the ears
 26. Pronasale—most anterior point of the nose at the midline
 27. Alare (right)—most lateral point of the nose
 28. Alare (left)—most lateral point of the nose
 29. Subalare—lowest point on the nose where the nose ridge meets the upper lip
 30. Labiale superius—highest midline point on the upper lip
 31. Cheilion (left)—most lateral point where the upper and lower lips meet
 32. Cheilion (right)—most lateral point where the upper and lower lips meet
 33. Stomion—imaginary point at the midline where the upper and lower lips meet
 34. Labiale inferius—lowest midline point on the lower lip
 35. Lowest midline point at the most anterior protrusion of the chin*
- Profile Points*
36. Vertex—highest point on the head, including hair
 37. Opisthocranion—most posterior point on the occipital bone
 38. Glabella—midline point between the eyebrows
 39. Superaurale—highest point on the ear
 40. Preaurale—point where the free portion of the auricle meets the skull
 41. Postaurale—most posterior point on the ear
 42. Porion—lowest point of the free auricle inside the ear
 43. Subaurale—lowest point of the earlobe
 44. Nasion—most indented point on the nose ridge within the orbital region
 45. Most anterior point on the pupil*
 46. Pronasale—most anterior point of the nose at the midline
 47. Columella—most anterior point of the nostril opening
 48. Alar curvature point—most posterior point of the nose curvature
 49. Subnasale—base of the nose where it meets the upper lip
 50. Labiale superius—highest midline point on the upper lip

51. Stomion—imaginary point at the midline where the upper and lower lips meet
52. Cheilion—most lateral point where the upper and lower lips meet
53. Labiale inferius—lowest midline point on the lower lip
54. Pogonion—most anterior point on the chin
55. Gnathion—lowest point on the chin

Three-Quarter View Points

56. Vertex—highest point on the head, including hair
57. Most posterior visual point of occipital bone*
58. Superciliare—highest point of each eyebrow
59. Nasion—most indented point on the nose ridge within the orbital
60. Superaurale—highest point on the ear
61. Preaurale—point where the free portion of the auricle meets the skull
62. Porion—lowest point of the free auricle inside the ear
63. Postaurale—most posterior point on the ear
64. Subaurale—lowest point of the earlobe
65. Frontosupraorbitale—point where brow ridge meets bony ridge under the eye
66. Exocanthion—lateral hinge where the eyelid closes
67. Palpebrale superius—highest point of the eyelid line when eye is relaxed open
68. Center point of the pupil*
69. Palpebrale inferius—lowest point of the upper eyelid line
70. Endocanthion—medial hinge where the eyelid closes
71. Orbitalis—bony ridge below the eye
72. Alare—most lateral point of the nose
73. Columella—most anterior point of the nostril opening
74. Pronasale—most anterior point of the nose at the midline
75. Cheilion—most lateral point where the upper and lower lips meet
76. Labiale superius—highest midline point on the upper lip
77. Lowest midline point at the most anterior protrusion of the chin*
78. Most visually posterior/basal point of the occipital bone*