Image and Vision Computing 29 (2011) 329–334

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Discussion

Reverse caricatures effects on three-dimensional facial reconstructions ☆

Jobany Rodriguez, Ricardo Gutierrez-Osuna *

Department of Computer Science and Engineering, Texas A&M University, TAMU 3112, College Station, TX 77843-3112, United States

ARTICLE INFO

Article history:
Received 9 April 2010
Received in revised form 30 October 2010
Accepted 3 January 2011

Keywords:
Face recognition
Face perception
Face reconstruction

ABSTRACT

Previous research has shown that familiarization with three-dimensional (3D) caricatures can help improve recognition of same-race and other-race faces, a result that may lead to new training tools in security applications. Since 3D facial scans are not generally available, here we sought to determine whether 3D reconstructions from 2D frontal images could be used for the same purpose. Our results suggest that, despite the high level of photographic realism achieved by current 3D facial reconstruction methods, additional research is needed in order to reduce reconstruction errors and capture the distinctive facial traits of an individual.

1. Introduction

We recognize faces from our own race better than those from another race [1–4]. This *other-race effect* (ORE) is robust, and has been reproduced in many studies [see 4–6] and across racial groups [6–8]. It is generally agreed that OREs result from the fact that the most appropriate features for discriminating faces are race-dependent [9]. For instance, Africans focus more on the shape and location of the eyes, eyebrows and ears, whereas Caucasians focus more on hair texture, and hair and eye color [10]. Research shows that this ORE may be reduced by drawing attention to the most distinctive feature of a given face. For example, Hills and Lewis [11] showed that OREs could be reduced by familiarizing subjects with own-race faces containing features critical for differentiating other-race faces.

At the same time, people have better recollection for visually distinctive faces [12–17], an effect that can be harnessed to create more memorable stimuli. Researchers pursuing this strategy create "caricatures" of normal faces by exaggerating their distinct qualities, and they find that people are more able to recognize these distorted faces than the veridical faces that were used to create them. This perceptual result is known as the *caricature effect* [18–21]. Additional studies have also demonstrated a *reverse-caricature effect*, according to which familiarization with caricatures improves the recognition of the veridical face at a later time [e.g., 18–24]. These results suggest ways in which caricatures may be used as training tools in applied face recognition settings (i.e. law enforcement[1]).

Motivated by this research, our previous work [27] has explored the use of three-dimensional (3D) caricatures to direct attention to critical features of other-race faces. Our experimental results showed that reverse-caricatures reduce OREs in Caucasian participants when viewing Indian faces. Although these results are a step towards designing real-life training systems, obtaining 3D models of individuals is cost prohibitive if not impossible in some applied settings; 3D scanners are still expensive instruments, and scanning is not possible if the target individual (i.e. a crime suspect) is at large. One potential solution to this problem is to use photogrammetric techniques to reconstruct 3D face models from 2D photographs [28,29]. Using these reconstructed 3D models one could then generate caricatures from individual mug shots. However, it is unclear whether caricatures based on reconstructed 3D models are still effective, since the caricaturization process may amplify reconstruction errors to the point of rendering the caricatures unusable. Answering this question is the main objective of this work.

2. Facial reconstruction and caricaturization

For this study, we used the University of Freiburg 3DFS-100 dataset [28] containing m = 100 3D face models. Each face consisted of a mesh with n = 75,972 vertices in full correspondence, and each vertex was defined by its position in 3D Cartesian coordinates $S=(X_1,Y_1,Z_1,X_2,\ldots,Y_n,Z_n)$ and its texture in RGB space $T=(R_1,G_1,B_1,R_2,\ldots,G_n,B_n)$. Performing principal component analysis [30], a face shape and texture can be defined by:

$$S = s_{avg} + \sum_i^{m-1} \alpha_i \cdot s_i \tag{1}$$

$$T = t_{avg} + \sum_i^{m-1} \beta_i \cdot t_i \tag{2}$$

☆ This paper has been recommended for acceptance by Lijun Yin.

* Corresponding author. Tel.: +1 979 845 2942; fax: +1 979 847 8578.
E-mail addresses: jobany@cse.tamu.edu (J. Rodriguez), rgutier@cse.tamu.edu (R. Gutierrez-Osuna).

[1] Caricatures also have been studied as a mechanism to improve recognition of criminal facial sketches or composites [25,26].

0262-8856/$ – see front matter
doi:10.1016/j.imavis.2011.01.002

where s_{avg} and t_{avg} are the shape and texture averages, α_i are the shape principal components, s_i are the shape eigenvectors, β_i are the texture principal components and t_i are the texture eigenvectors; examples for same-race and other-race faces from the dataset are illustrated in Fig. 1a.

To test whether 3D reconstructions are amenable to reverse-caricature effects, we decided to use a best-case reconstruction scenario. Namely, we reconstructed the shape and texture of each face in the 3DFS-100 dataset in a leave-one-out fashion while holding constant the rendering parameters (i.e., camera position and illumination). First, we removed each face from the dataset and obtained a PCA decomposition on the remaining m = 99 training faces according to their shape (S_{train}) and texture (T_{train}). Next, we projected the left-out test face f_{test} along the PCA eigenvectors (s_i, t_i) to obtain its principal components α_{test} and β_{test}:

$$\alpha_{test} = \left(S_{test} - s_{avg}\right)^{\mathrm{T}} \cdot s_i \tag{3}$$

$$\beta_{test} = \left(T_{test} - t_{avg}\right)^{\mathrm{T}} \cdot t_i \tag{4}$$

which yield a reconstructed 3D model f'_{test} :

$$S'_{test} = s_{avg} + \sum_i^{m-1} \alpha_{test} \cdot s_i \tag{5}$$

$$T'_{test} = t_{avg} + \sum_i^{m-1} \beta_{test} \cdot t_i \tag{6}$$

Fig. 1. Sample stimuli: (a) ground-truth frontal faces, (b) their corresponding 3D segment-based reconstructions, (c) caricatures from ground-truth faces, and (d) caricatures from reconstructed faces. Ears and neck were manually removed to prevent participants from using picture-matching strategies. Inspection of (c) and (d) illustrates the extent to which caricatures amplify reconstruction errors rather than unique facial traits.

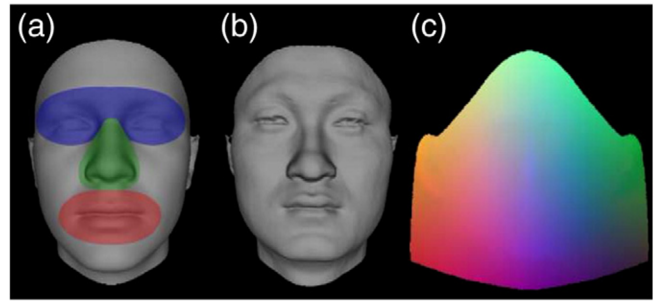

Fig. 2. (a) Face segmentation used for 3D face reconstructions; each region is predicted independently and then merged into a composite face. (b) Example of a 3D shape and (c) its corresponding geometry image; the geometry image is an $n \times m$ matrix where XYZ coordinates are represented as RGB values.

2.1. Blending segment-wise reconstructions

Face reconstructions f'_{test} have $2(m-1)$ degrees of freedom ($m-1$ associated with shape, and $m-1$ associated with texture). To increase the level of expressiveness, we segmented the face into four regions [28], and performed the PCA decomposition for each segment independently; see Fig. 2a. The final face model f'_{test} was obtained by combining each predicted segment through an image blending procedure [31]. Namely, given two input images (A and B) to be blended, we define a mask image M per segment that denotes whether the corresponding pixel should come from image A ($M_{ij} = 1$) or B ($M_{ij} = 0$). Then we construct a Laplace pyramid for images A and B, and a Gaussian pyramid for the mask image M, as illustrated in Fig. 3. At each level in the pyramid, the algorithm blends the two images as:

$$LC_n(i,j) = GM_n(i,j) \cdot LA_n(i,j) + (1 - GM_n(i,j)) \cdot LB_n(i,j) \tag{7}$$

where LA_n, LB_n, and LC_n are the Laplace pyramids of the input images (A and B,) and the output image C, respectively, and GM_n is the Gaussian pyramid of the mask image for a given level n. Finally, the resulting blended image C , is synthesized from the LC_n pyramid as:

$$G_n = LC_n \tag{8}$$

$$G_{n-1} = LC_{n-1} + \mathrm{Expand}(G_n) \tag{9}$$

In order to apply this image-based blending algorithm to 3D models, the 3D segments are converted into geometry images (GI) [32] prior to the blending stage; see Fig. 2b,c. After blending all GI-based segments, the resulting GI is converted back into a 3D model. The overall method produces seamless and photorealistic reconstructions that are comparable to those in previous work [28,29,33]; examples for same-race and other-race faces are illustrated in Fig. 1b.

2.2. Caricaturization

In order to caricaturize faces consistently and evenly,[2] we first normalize each face f by its Mahalanobis distance ($\|\ \|_{\mathrm{M}}$) to the average face f_{AVG} [35]:

$$f_N = \frac{f}{\|f - f_{AVG}\|_M} \tag{10}$$

and then caricaturize it by linearly exaggerating differences with respect to f_{AVG} [19,21,35]:

$$f_C = f_{AVG} + (1 + \alpha)(f_N - f_{AVG}) \tag{11}$$

[2] Distinctive faces need to be caricaturized less than typical faces in order to achieve the same level of distinctiveness [34].

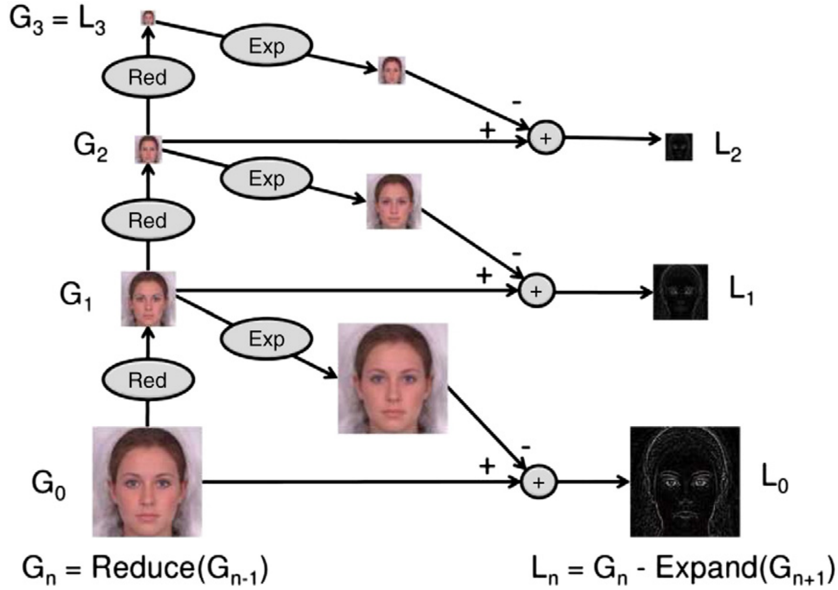

Fig. 3. Construction of the Laplace pyramid for an image. The process starts with a Reduce () operation (low-pass filtering) and sub-sampling (half sized) the original image, G_0, to obtain G_1. This process continues until reaching a predefined level N. G_N is known as the Gaussian pyramid. To complete the Laplace pyramid L_N, a band-pass operation is required between two successive low-pass levels. The lower frequency image (G_3 in the example), is interpolated using the Expand() operation before subtracting it from the higher frequency image, G_2. This process continues until reaching L_0.
Adapted from [31].

where the average face f_{AVG} is computed using the 99 training faces. We define caricaturization levels (α) as a function of the standard deviation (σ) of the distance between un-normalized faces in the dataset and their average:

$$\sigma = std(\|f - f_{AVG}\|_M) \tag{12}$$

This parameterization is preferable to the conventional percentage factor (α) because it adjusts the caricaturization level to the intrinsic variability of faces in the dataset; as an example, a caricature factor $\alpha = 10\%$ may be excessive for a fairly homogeneous dataset or insignificant for a very diverse dataset, whereas a caricature factor of $\sigma = 1$ accounts for the variability in the dataset. Note that the relationship between σ and α is trivially defined by:

$$\alpha = \frac{k\sigma}{\|f_N - f_{AVG}\|_M}; \forall k \in \mathbb{Z} \text{ s.t. } \alpha > -1 \tag{13}$$

Fig. 1c shows the results of applying this caricaturization method to the ground-truth faces in Fig. 1a, whereas Fig. 1d shows the equivalent caricatures when obtained from the reconstructed faces in Fig. 1b. The reconstruction method is able to capture the main characteristics of a face (e.g., gender, race, anthropometric cues), but also introduces a few reconstruction artifacts that become evident as we compare the two caricature sets. Though more noticeable in other-race faces, reconstruction artifacts also occur in Caucasian faces; e.g., the reconstruction fails to capture the unique chin dimple in the second face in Fig. 1b. To what extend do these reconstruction errors hamper facial recognition by humans?

3. Perceptual experiments

To answer this question, we employed an old/new face recognition protocol [36,37] whereby subjects are first familiarized with a set of faces, and then asked to recognize those faces among a set of confounders. Following [27], we used the minimal degree of exaggeration between veridical and caricature faces that would lead to a caricature effect. This exaggeration level was then used in a recognition study that allowed us to test whether familiarization with caricatures of *reconstructed* 3D models (as opposed to caricatures of the *original* 3D models) would reduce OREs.

3.1. Stimuli

Forty face models were selected from the 3DFS-100 dataset [28]. The same 40 faces were used throughout the experiments. From these ground-truth veridical faces (V) we generated 40 veridical reconstructions (Vr), which were in turn used to generate 40 caricaturized reconstructions (Cr). Following Furl et al. [37], our face corpus had a similar distribution across races: 10 Caucasians, 10 East Asians, and 10 Indians. We also included 6 African faces and 4 faces from other groups (Middle Eastern, Hispanic) as filler stimuli. East Asian and Indian faces were treated as other-race faces.

3.2. Procedure

Forty-three Caucasian undergraduate students (24 females and 19 males) from the Department of Psychology at Texas A&M University participated in this study. Participants were assigned to one of two experimental conditions:

- Vr-V condition: familiarization with veridical reconstructions (Vr), recognition of veridical faces (V). Twenty-one students participated in this study, which served as a control.
- Cr-V condition: familiarization with caricaturized reconstructions (Cr), recognition of veridical faces (V). Twenty-two students participated in this condition, which tested our working hypothesis.

For each condition, participants were familiarized with 20 frontal target faces (the same faces for all subjects), each presented twice in random order, 3 seconds per presentation. Following familiarization, participants were tested on 40 faces, of which 20 were "new" (non-target) and 20 were "old" (target); all faces in the test phase were rendered with a random orientation between ±5 degrees in the three axes in order to prevent picture-matching strategies [24]. Participants were asked to identify each face as "old" if they recognized it as one from the familiarization phase, or as "new" otherwise. Following [36], no time limits were imposed, but participants were asked to make the

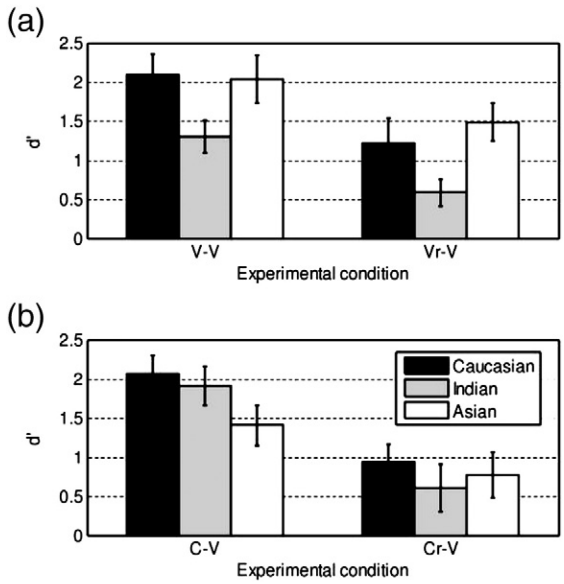

Fig. 4. Signal detection d' for (a) V-V and Vr-V conditions, and (b) C-V and Cr-V conditions. V-V and C-V conditions are based on [27] results. Error bars represent standard errors.

identification as rapidly as possible without sacrificing accuracy. In all conditions, the same 20 randomly chosen faces were used (in either caricature or veridical form). Following [27] exaggeration levels were set to -1σ ($\alpha = -0.21$) and 0σ ($\alpha = 0$) for veridical and caricaturized faces, respectively. A similar gender distribution for participants was maintained across experimental conditions.

3.3. Results

3.3.1. Do 3D reconstruction errors affect recognition performance?

To answer this question, we compared results from the Vr-V condition (i.e., in the new experiments above) against those on the V-V condition (i.e., in our earlier study [27][3]) in terms of the signal detection d' measure [38]. Results are summarized in Fig. 4 Using the experimental condition (V-V vs. Vr-V) and race (Caucasian vs. Asian vs. Indian) as factors, a two-way ANOVA shows a main effect of experimental condition, $F(2,82) = 8.735$, $p < 0.01$. Overall performance in the Vr-V condition is lower than in the V-V condition: $t(42) = 2.956$, $p < 0.01$, and also on each individual race. Thus, these results indicate that reconstruction errors have a negative effect on recognition performance, regardless of race.

To tease apart the influence of reconstruction errors on different decisions, we also analyzed results in terms of true positive rates (TPR) and false positive rates (FPR). Results are summarized in Fig. 5:

- TPRs showed a main effect of race $F(2,84) = 6.95$, $p < 0.01$, and of experimental condition, $F(1,42) = 18.108$, $p < 0.001$. There was no interaction effect between race and experimental condition, $F(2,84) = 0.444$, *ns*. Overall performance on the V-V condition was significant higher than on the Vr-V condition, $t(42) = 4.255$, $p < 0.001$.
- FPRs showed a main effect of race $F(2,84) = 4.99$, $p < 0.05$, and interaction effects between race and experimental condition, $F(2,84) = 3.125$, $p < 0.05$. There is no main effect of experimental condition, $F(1,42) = 0.687$, $p > 0.05$. Overall performance on the V-V condition was not significant different than on the Vr-V condition, $t(42) = 0.829$, $p > 0.05$.

[3] Our earlier experiments used the same procedure, stimulus set, and subject pool as those reported here.

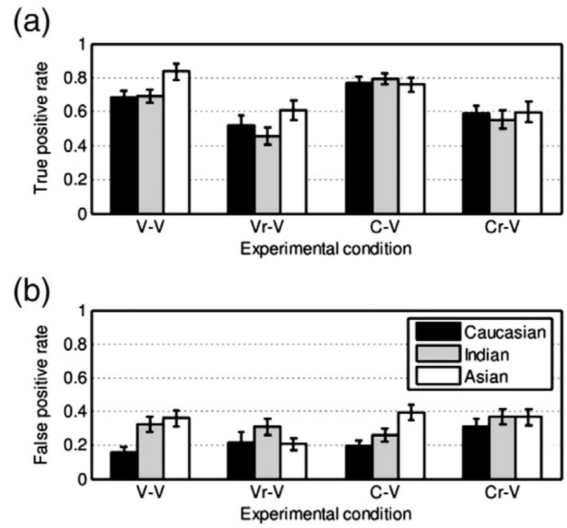

Fig. 5. (a) True positive and (b) false positive rates across experimental conditions. V-V and C-V conditions are based on [27] results. Error bars represent standard errors.

These results indicate that reconstruction errors affected recognition performance by decreasing TPRs but not necessarily by increasing FPRs. However, there was a main effect of race on both measures, which suggests that reconstruction errors affect the viewer's performance differently depending on the race of the stimulus face.

Finally, we analyzed the signal detection criterion C, which provides cues about whether participants have a bias toward a particular answer. A conservative participant will answer 'no' more often (positive C values), while a liberal participant will respond 'yes' more often (negative C values) [39]. Results are shown in Fig. 6a. Analysis of variance shows a main effect of race: $F(2,84) = 7.55$, $p < 0.001$ and a main effect of experimental condition: $F(1,42) = 9.621$, $p < 0.01$, but no interaction effects: $F(2,84) = 2.914$, *ns*. Participants on the V-V condition were more conservative to own-race faces than to other-race faces. In addition, training on the Vr-V condition caused a noted increase in criterion C, which indicates that reconstruction errors made participants more conservative. We can infer participants were

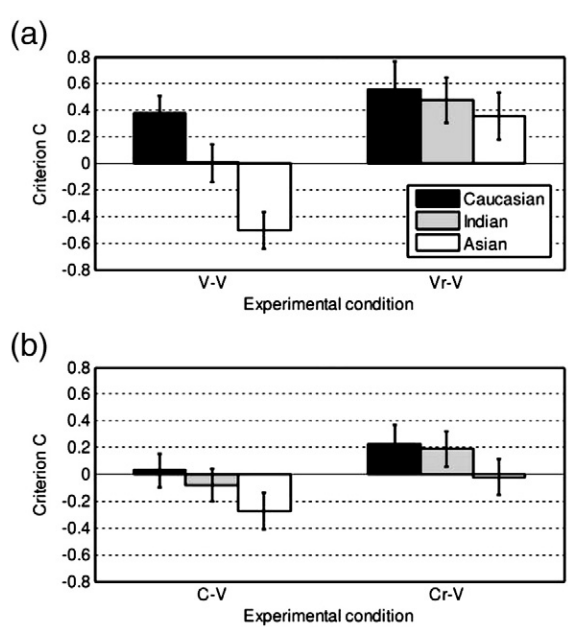

Fig. 6. Criterion C across experimental conditions. V-V and C-V conditions are based on [27] results. Error bars represent standard errors.

having difficulties learning from reconstructions. This is consistent with the TPRs results above; conservative participants would experience a reduced TPR score. Furthermore, conservative participants would improve their FPRs. In this pair of experimental conditions, participants maintained the same level of FPR scores.

3.3.2. Do caricatures reduce the impact of 3D reconstruction errors on recognition performance?

To answer this question, we compared results in the Vr-V and Cr-V conditions, also in terms of the signal detection d' measure. Using experimental condition (Vr-V vs. Cr-V) and race (Caucasian vs. Asian vs. Indian) as factors, a two-way ANOVA shows no main effect of experimental condition, $F(1,41)=2.068$, $p>0.05$. However, performance was lower on Cr-V than on Vr-V: $t(41)=1.438$, *ns*, ($p=0.07$; one-tailed). There was no main effect of race, $F(2,82)=2.519$, *ns*. OREs on Indian faces were observed for participants in the Vr-V condition: $t(40)=1.69$, $p<0.05$ (one-tailed), but not for participants in the Cr-V condition: $t(42)=0.88$, *ns*. However, this reduction of OREs with reverse caricatures was due to a reduction in recognition performance on Caucasian faces rather than to an increase in recognition performance on Indian faces. There was no significant interaction between race and training condition, $F(2,82)=0.997$, *ns*; both training conditions show similar performance profile across races. Therefore, when considering the signal detection d' measure, reverse-caricature training using reconstructed 3D models did not reduce the impact of reconstruction errors. In fact, it appears that the caricature process amplifies reconstruction errors to a greater extent that it makes the unique facial traits more salient.

As before, we performed a finer-grained analysis in terms of true positive rates (TPR) and false positive rates (FPR). Results are summarized in Fig. 5a-b.

- TPRs did not show a main effect of race $F(2,82)=2.096$, *ns*, training condition, $F(1,41)=0.959$, *ns*, or interaction effects, $F(2,82)=0.681$, *ns*. However, performance was better on Cr-V than it was on Vr-V: $t(41)=0.972$, $p>0.05$.
- FPRs did not have a main effect of race: $F(2,82)=1.275$, *ns*, but had a significant main effect of condition: $F(1,41)=6.332$, $p<0.05$. Namely, performance was worse on Cr-V than on Vr-V: $t(41)=2.516$, $p<0.05$.

These results suggest that although the caricature process may improve TPRs, this is at the expense of a much larger increase in FPRs such that the net effect of caricaturization (as measured by the signal detection d' measure) is detrimental.

Finally, we also analyzed the signal detection criterion C. Results are shown in Fig. 6. Analysis of variance shows no main effect of race, $F(2,82)=1.424$, *ns*, and no significant interaction, $F(2,82)=0.050$, *ns*. However, there is a main effect of training condition $F(1,41)=3.80$, $p<0.05$. Namely, participants in the Cr-V condition become less conservative than those in the Vr-V condition and approach the ideal observer (i.e., $C=0$, no strategy).

3.3.3. Do caricatures of reconstructed faces provide better recognition performance than veridical faces?

To answer this question, we compared results in the Cr-V and V-V conditions in terms of the signal detection d' measure. Using experimental condition (Cr-V vs. V-V) and race (Caucasian vs. Asian vs. Indian) as factors, a two-way ANOVA did not show a main effect of race, $F(2,86)=2.796$, *ns*, and did not have an interaction between race and experimental condition. There was a main effect of experimental condition, $F(1,43)=17.792$, $p<0.001$; performance on the V-V condition was significantly higher than on the Cr-V condition, $t(43)=4.218$, $p<0.001$. Thus, the reverse-caricature effect we observed in our earlier work [27] disappears when caricatures are obtained from reconstructed 3D models, which suggests that reconstruction errors can perceptually mask distinctive facial cues.

As before, we also analyzed TPRs and FPRs; results are summarized in Fig. 5:

- TPRs did not show a main effect of race $F(2,86)=2.512$, *ns*, and no interaction effects, $F(2,86)=1.313$, *ns*. TPRs showed a main effect of training condition, $F(1,43)=12.09$, $p<0.001$. Namely, Cr-V performance was significantly lower than V-V performance: $t(43)=3.477$, $p<0.001$.
- FPRs had a main effect of race, $F(2,86)=5.501$, $p<0.01$, but no effect of condition, $F(1,43)=3.184$, *ns*, or interaction effects: $F(2,86)=1.759$, *ns*.

These results mirror those in Section 3.3.1 and indicate that the lower recognition performance on Cr-V is due to a reduction of TPRs rather than an increase in FPRs.

Finally, analysis of variance on the signal detection criterion C, summarized in Fig. 6, shows no main effect of experimental condition: $F(1,43)=2.204$, *ns*, or interaction effects: $F(2,86)=2.799$, *ns*. However, there is a main effect of race: $F(2,86)=9.031$, $p<0.001$, namely in terms of reduced OREs in the Cr-V condition.

4. Discussion

By definition, caricatures increase the salience of idiosyncratic or normatively distinct qualities. This form of distortion appears to increase the amount of memorable information available for later recognition. As a result, exposure to facial caricatures can increase later recognition of their veridical counterparts and also can reduce the ORE. Our previous study [27] showed that caricatures from ground-truth 3D models improve the recognition of their veridical counterparts and also reduce OREs, as measured by the signal detection's sensitivity index d' [38]. The main objective of this work was to determine whether these earlier results would extend to face recognition when the veridical 3D models are not available and have to be replaced by reconstructions.

Our results indicate that 3D reconstructions are not of sufficient quality to be used for face recognition purposes, even when rendered without caricaturization. Training with reconstructed faces (Vr-V) leads to lower TPRs when compared to training on veridical faces (V-V), although FPRs seem immune to reconstruction errors. Training on caricatures of reconstructed faces (Cr-V) leads to higher TPRs when compared to training on reconstructed faces (Vr-V) but at the expense of a larger increase in FPRs, with a negative net effect. This suggests that the caricature process amplifies reconstruction errors more than it enhances distinctive facial features. Finally, training with caricatures of reconstructions (CrV) leads to lower recognition performance than than training on veridical faces (V-V), mainly in terms of reduced TPRs.

Collectively, our study indicates that the reconstruction process fails to capture the more distinctive features of a given face (e.g., notice the missing chin dimple in the second row of Fig. 1.). Because caricatures amplify differences relative to a norm, they also exacerbate any errors introduced during reconstruction (e.g., first row face in Fig. 1), possibly distracting participants away from the distinctive features of faces. Thus, our results suggest that the original 3D faces may be required in order to generate perceptually valid caricatures.

These results are critical for the development of training tools because in most realistic settings 3D scans of the target faces are not available. In these cases, all that is available is a 2D “mug shot” of the target. While at the onset of this study we anticipated that caricatures would amplify reconstruction errors, it was not clear whether these errors would compromise human's ability to recognize faces. As an example, humans recognize faces under fairly severe manipulations (e.g., at very low resolution, under partial occlusion), but this ability breaks down with other types of lossless manipulations (e.g., rotated or inverted faces); see [40].

4.1. Future work

One may be tempted to infer that more sophisticated reconstruction algorithms would be needed, such as the non-linear optimization method in [28]. While this may be the case in more general reconstruction scenarios (e.g., with real photographs), the reconstruction method in Eqs. (5) and (6) is optimal (in the mean-square-error sense) for our study since the representation is linear in the optimization parameters (shape and texture) and all remaining parameters (camera and illumination) were assumed known. Thus, failure to reproduce our earlier results (which used ground-truth 3D faces) must be attributed to factors other than the reconstruction procedure, and instead point in the direction of facial databases. Specifically, extending the 3D face database would certainly improve the reconstruction quality, as additional faces would provide a more diverse pool from which to reconstruct new faces. Likewise, increasing the number of input images of a probe face (e.g. frontal, 3/4 and profile views) might help capture distinctive facial traits that are not prominent in a frontal view. Facial diversity (race, gender, age) most likely plays a significant role in the quality of reconstruction. Having a well-balanced database or one that specifically matches the characteristics of the input face (e.g. Caucasian, male, 55-60 years of age) might also improve reconstruction results. Additional research is also needed to determine, among others, (*i*) the level of reconstruction accuracy that must be achieved to obtain *perceptually* realistic (as opposed to photorealistic) results, (*ii*) the type and number of facial segments (Fig. 2a), maybe on a race-by-race basis or to account for facial asymmetries, and (*iii*) the types of reconstruction errors (e.g. shape vs. reflectance, different facial areas) that have the greatest impact on recognition performance. Addressing these questions is a necessary step towards the development of effective training tools for face recognition.

Our work evaluates whether reverse-caricatures can improve the recognition of *specific* other-race faces, i.e. a closed-world assumption. While we believe that, with sufficient exposure to caricaturized other-race faces, reduction of OREs may generalize to faces not seen in the familiarization phase (see [11]), this issue will also require further investigation. However, in security applications where one seeks to improve the recognition of specific faces (i.e. those of a suspect at large), the closed-world assumption is valid.

Acknowledgment

We are grateful to Hart Blanton and Heather Bortfeld in the Department of Psychology at the University of Connecticut for assistance with the perceptual experiments. This work was funded in part by TAMU-CONACYT contract #246641.

References

[1] T. Valentine, M. Endo, Towards an Exemplar Model of Face Processing: The Effects of Race and Distinctiveness, The Quarterly Journal of Experimental Psychology 44 (1992) 671–703.
[2] D.S. Lindsay, P.C. Jack, M.A. Christian, Other-Race Face Perception, Journal of Applied Psychology 76 (1991) 587–589.
[3] T. Valentine, V. Bruce, The effect of race, inversion and encoding activity upon face recognition, Acta Psychologica 61 (1986) 259–273.
[4] C.A. Meissner, J.C. Brigham, Thirty Years of Investigating the Own-Race Bias in Memory for Faces: A Meta-Analytic Review, Psychology, Public Policy, and Law 7 (2001) 3–35.
[5] P. Shapiro, S.D. Penrod, A meta analysis of the facial identification literature, Psychological Bulletin 100 (1986) 139–156.
[6] R.K. Bothwell, J.C. Brigham, R.S. Malpass, Cross-Racial Identification, Personality and Social Psychology Bolletin 15 (1989) 19–25.
[7] S.M. Kassin, P.C. Ellsworth, V.L. Smith, On the "General Acceptance" of Eyewitness Testimony Research: A New Survey of the Experts, American Psychologist 44 (1989) 1089–1098.
[8] J.C. Brigham, L.B. Bennett, C.A. Meissner, T.L. Mitchell, The influence of race on eyewitness memory, in: R. Lindsay, D. Ross, J. Read, M. Toglia (Eds.), Handbook of Eyewitness Psychology: Memory for People, Lawrence Erlbaum & Associates, 2007.
[9] J. Shepherd, J. Deregowski, Races and faces - a comparison of the responses of Africans and Europeans to faces of the same and different races, British Journal of Social Psychology 20 (1981) 125–133.
[10] H. Ellis, J. Deregowski, J. Shepherd, Description of white and black faces by white and black subjects, International Journal of Psychology 10 (1975) 119–123.
[11] P.J. Hills, M.B. Lewis, Reducing the own-race bias in face recognition by shifting attention, The Quarterly Journal of Experimental Psychology 59 (2006) 996–1002.
[12] M. Going, J.D. Read, The effect of uniqueness, sex of subject, and sex of photograph on facial recognition, Perceptual and Motor Skills 39 (1974) 109–110.
[13] G.L. Wells, A. Memon, S. Penrod, Eyewitness evidence: Improving its probative value, Psychological Science in the Public Interest 7 (2006) 45–75.
[14] T. Valentine, V. Bruce, The effects of distinctiveness in recognising and classifying faces, Perception 15 (1986) 525–535.
[15] L. Light, F. Kayra-Stuart, S. Hollander, Recognition memory for typical and unusual Faces, Journal of Experimental Psychology: Human Learning and Memory 5 (1979) 212–228.
[16] T. Valentine, A unified account of the effects of distinctiveness, inversion and race in face recognition, Quarterly Journal of Experimental Psychology 43A (1991) 161–204.
[17] T. Valentine, V. Bruce, Recognizing Familiar Faces: The Role of Distinctiveness and Familiarity, Canadian Journal of Psychology 40 (1986) 300–305.
[18] P.J. Benson, D.I. Perrett, Visual processing of facial distinctiveness, Perception 23 (1994) 75–93.
[19] K. Lee, G. Byatt, G. Rhodes, Caricature effects, distinctiveness, and identification: Testing the face-space framework, Psychological Science 11 (2000) 379–385.
[20] R. Mauro, M. Kubovy, Caricature and face recognition, Memory & Cognition 20 (1992) 433–440.
[21] G. Rhodes, S. Brennan, S. Carey, Identification and ratings of caricatures: implications for mental representation of faces, Cognitive Psychology 19 (1987) 473–497.
[22] K.A. Deffenbacher, J. Johanson, T. Vetter, A.J. O'Toole, The face typicality-recognizability relationship: Encoding or retrieval locus? Memory & Cognition 28 (2000) 1173–1182.
[23] S.V. Stevenage, Can caricatures really produce distinctiveness effects? British Journal of Psychology 86 (1995) 127–146.
[24] J. Rodríguez, H. Bortfeld, I. Rudomín, B. Hernández, R. Gutiérrez-Osuna, The Reverse-caricature Effect Revisited: Familiarization with Frontal Facial Caricatures Improves Veridical Face Recognition, Applied Cognitive Psychology 23 (2009) 733–742.
[25] C.D. Frowd, V. Bruce, D. Ross, A. McIntyre, P.J.B. Hancock, An application of caricature: how to improve the recognition of facial composites, Visual Cognition 15 (2007) 954–984.
[26] Y. Zhang, C. McCullough, J.R. Sullins, C.R. Ross, Hand-drawn Face Sketch Recognition by Humans and A PCA-based Algorithm for Forensic Applications, IEEE Transactions on Systems, Man, and Cybernetics Part A 40 (2010) 475–485.
[27] J. Rodríguez, H. Bortfeld, R. Gutiérrez-Osuna, Reducing the Other-Race Effect through Caricatures, 8th IEEE International Conference on Automatic Face and Gesture Recognition, Amsterdam, NL, 2008.
[28] V. Blanz, T. Vetter, A morphable model for the synthesis of 3D faces, 26th annual conference on computer graphics and interactive techniques, ACM Press/ Addison-Wesley Publishing Co, 1999, pp. 187–194.
[29] Y. Hu, D. Jiang, S. Yan, L. Zhang, H. Zhang, Automatic 3D reconstruction for face recognition, Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004, pp. 843–848.
[30] S. Roweis, Finding the first few eigenvectors in a large space, 1996.
[31] P.J. Burt, E.H. Adelson, Merging Images through Pattern Decomposition, Proc. Applications of Digital Image Processing VIII, 1985, pp. 187–194.
[32] X. Gu, S. Gortler, H. Hoppe, Geometry images, 29th annual conference on Computer graphics and interactive techniques, ACM Press, San Antonio, Texas, 2002, pp. 355–361.
[33] V. Blanz, T. Vetter, Face Recognition Based on 3D Shape Estimation from Single Images, CGF-TR (2002) 2.
[34] G. Rhodes, I.G. McLean, Distinctiveness and expertise effects with homogeneous stimuli: Towards a model of configural coding, Perception 19 (1990) 773–794.
[35] A.J. O'Toole, T. Vetter, H. Volz, E.M. Salter, Three-dimensional caricatures of human heads: distinctiveness and the perception of facial age, Perception 26 (1997) 719–732.
[36] B. Duchaine, T. Nieminen-von Wendt, J. New, T. Kulomaki, Dissociations of visual recognition in a developmental prosopagnosic: Evidence for separate developmental processes, Neurocase 9 (2003) 380–389.
[37] N. Furl, P.J. Phillips, A.J. O'Toole, Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis, Cognitive Science 26 (2002) 797–815.
[38] N.A. Macmillan, C.D. Creelman, Detection Theory: A User's Guidesecond ed, , 2005.
[39] H. Abdi, Signal detection theory, in: Salkind (Ed.), Encyclopedia of Measurements and Statistics, Thousand Oaks, N.J, 2007, pp. 886–889.
[40] P. Sinha, B. Balas, Y. Ostrovsky, R. Russell, Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About, Proceedings of the IEEE 94 (2006) 1948–1962.