Science and Justice xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Science and Justice

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

Recognition of computerized facial approximations by familiar assessors

Adam H. Richard [a], Keith L. Monson [b,*]

[a] Federal Bureau of Investigation, Laboratory Division, 4940 Fowler Road, Redstone Arsenal, Huntsville, AL 35898, United States
[b] Federal Bureau of Investigation, Laboratory Division, Counterterrorism and Forensic Science Research Unit, 2501 Investigation Parkway, Quantico, VA 22135, United States

ARTICLE INFO

Article history:
Received 22 February 2017
Received in revised form 13 June 2017
Accepted 18 June 2017
Available online xxxx

Keywords:
Facial reconstruction
Facial reproduction
Facial approximation
Facial recognition
Unidentified human remains
Missing persons

ABSTRACT

Studies testing the effectiveness of facial approximations typically involve groups of participants who are unfamiliar with the approximated individual(s). This limitation requires the use of photograph arrays including a picture of the subject for comparison to the facial approximation. While this practice is often necessary due to the difficulty in obtaining a group of assessors who are familiar with the approximated subject, it may not accurately simulate the thought process of the target audience (friends and family members) in comparing a mental image of the approximated subject to the facial approximation. As part of a larger process to evaluate the effectiveness and best implementation of the ReFace facial approximation software program, the rare opportunity arose to conduct a recognition study using assessors who were personally acquainted with the subjects of the approximations. ReFace facial approximations were generated based on preexisting medical scans, and co-workers of the scan donors were tested on whether they could accurately pick out the approximation of their colleague from arrays of facial approximations. Results from the study demonstrated an overall poor recognition performance (i.e., where a single choice within a pool is not enforced) for individuals who were familiar with the approximated subjects. Out of 220 recognition tests only 10.5% resulted in the assessor selecting the correct approximation (or correctly choosing not to make a selection when the array consisted only of foils), an outcome that was not significantly different from the 9% random chance rate. When allowed to select multiple approximations the assessors felt resembled the target individual, the overall sensitivity for ReFace approximations was 16.0% and the overall specificity was 81.8%. These results differ markedly from the results of a previous study using assessors who were unfamiliar with the approximated subjects. Some possible explanations for this disparity in performance were examined, and it was ultimately concluded that ReFace facial approximations may have limited effectiveness if used in the traditional way. However, some promising alternative uses are explored that may expand the utility of facial approximations for aiding in the identification of unknown human remains.

1. Introduction

Facial approximation is an investigative tool used to generate leads in cases of unidentified human remains. Methods of positive identification, including DNA, dental, and medical record comparisons, are not possible without some sort of antemortem record, or a putative blood relative, against which to compare. This requires that the person has been reported missing and/or that their records are available to investigators in some searchable format. If neither of these requirements is in place, one investigative option for identification is to reach out to the public in the hopes that someone will provide a possible identity that can be further investigated as a possible match. Facial approximations can be used to facilitate this public outreach by providing an estimation of what the unidentified person may have looked like in life. The goal is for friends or relatives of the missing person to recognize the facial approximation and contact authorities with details of their missing person's identity for verification or exclusion.

Facial approximations traditionally have been sculpted in clay by forensic artists. Facial approximation artists, ideally in collaboration with physical anthropologists, use the architecture of unidentified skulls, research on demographically-specific facial tissue depths, and the knowledge of craniofacial anatomy to create estimations of the facial appearance of unidentified individuals. In recent years, the FBI Laboratory, in conjunction with G.E. Global Research, has developed a software program to generate facial approximations digitally. Reality Enhancement Facial Approximation by Computational Estimation, or ReFace, is a program designed to use a CT or laser scan of an unidentified skull to create a statistically-based approximation of that unidentified person's appearance in life [1]. Because ReFace is still undergoing testing, it is not currently available to the public. However, over the past several years, multiple studies have tested the accuracy and efficacy of using ReFace approximations for potential application in the quest to identify the nation's growing number of unidentified human remains. Some of these studies examined quantitative similarities and differences between

* Corresponding author.
E-mail address: Keith.Monson@ic.fbi.gov (K.L. Monson).

http://dx.doi.org/10.1016/j.scijus.2017.06.004
1355-0306/

Please cite this article as: A.H. Richard, K.L. Monson, Recognition of computerized facial approximations by familiar assessors, Sci. Justice (2017), http://dx.doi.org/10.1016/j.scijus.2017.06.004

ReFace approximations and known facial topography from CT scan donors [2–5], while other studies have looked at the potential for human participants to recognize these approximations [6–8]. Automated facial recognition of ReFace approximations shows particular promise [2,4,5,9–12]. Thus far, ReFace human recognition studies have involved assessors who were unfamiliar with the facial approximation subjects [7,8]. The availability of medical scans from co-workers provided a rare opportunity to test the potential of human recognition of ReFace approximations in a situation as close to operationally real as is feasible, allowing for a comparison with results from previously published recognition studies involving ReFace.

As it is very difficult to create a test scenario involving assessors who are familiar with the approximation subject, only a few such studies have been conducted [13–16]. This is not ideal given that, operationally, the target audience for facial approximations is friends and loved ones of the missing person in question. Of the familiar assessor studies available in the literature, one study involved computer manipulation of 2D facial images to simulate a best-case scenario for facial approximation recognition [13]. A group of facial images from 130 females of European ancestry were combined to create an "average" face; this face was then applied to the facial outlines from frontal photos of five study participants to create facial approximations of those individuals. These approximations, while having an averaged color and texture, conformed to the exact facial shape of their targets, making them more accurate than would normally be expected for a forensic facial approximation. The manipulated facial images were shown to both familiar and unfamiliar assessors. In the case of familiar assessors, the participants were asked to indicate if they recognized each of the five facial images. In the unfamiliar assessor tests, a pool of six photographs was provided with each of the five facial manipulations. The unfamiliar assessor testing resulted in 43% true positive, 28% false positive, and 29% false negative rates (29% TP, relative to chance[1]) when the target face was present in the pool of photographs. True negative rates when no target was present in the pool were not statistically different from chance. The three target approximations used in the familiar assessor tests were recognized 86%, 40%, and 0% of the time, respectively, for an average recognition rate of 42% (28% above chance). In both the familiar and unfamiliar tests, siblings participated in the test for two of the subjects, but failed to recognize the individual, despite knowing the subject for 20 years. This value was proposed as the ceiling recognition limit one could expect to achieve using a 2D facial approximation given that the images in this study were more accurate than a typical forensic facial approximation.

A pair of related facial approximation studies involving familiar assessors was conducted by Fernandes et al. [17,18]. The first of these studies involved the creation of computer-generated facial approximations using cranial data from a CT scan and three different facial tissue depth standards [17]. One facial approximation was made using each of the three tissue depth standards. The same hair style, eye brows, and eyelashes were applied to each approximation, and then individuals who knew the donor were shown an approximation and ten pictures of individuals they were familiar with and were required to indicate to whom they believed the approximation belonged. Results from the study found recognition rates ranging from 20 to 27% (10–17% above chance), depending on which tissue depth standard had been used.

In the second study, the same three approximations were used, but this time without the addition of hair [18]. Again, individuals who were familiar with the facial approximation subject were shown the approximation along with an array of ten faces belonging to people they knew and were told to select whom they thought the approximation represented. Interestingly, the recognition rates for approximations lacking hair ranged from 32 to 41% (22–31% above chance), consistently higher than any of the approximations having hair in the previous study. They concluded that including hair in a facial approximation, especially when no evidence for the unidentified individual's hair color or style is present, can be detrimental to the recognition potential of the approximation. This is also interesting given that the approximations from Stephan et al. [13] were all presented with the same short, brown hair, and the approximation that resulted in no correct identifications by familiar assessors was the only blonde target in the group. The other two targets both had brown hair, although the styles were different from that of the approximated faces.

2. Methods

2.1. Facial approximation creation

ReFace is designed to use laser or computed tomography (CT) scans of skulls from unidentified human remains to generate approximations of what the unidentified individual may have looked like in life. Facial approximations for this study were generated from previously acquired CT or magnetic resonance imaging (MRI) scans (use of the latter is discussed below) donated by eight volunteers (Table 1) [redacted] who also provided basic demographic data (sex, age, and ancestry). A ninth individual (Donor 2) who lacked medical head scans volunteered to be a negative control for the study. An array of ten foil approximations was created for the negative control to see if participants would feel compelled to make a selection, even in the absence of a correct choice. Human subject involvement was approved by the [redacted] Institutional Review Board and each donor provided informed consent. Cranial data were segmented from the medical head scans of the donor using Mimics v17 (Materialise, Leuven, Belgium), then imported into ReFace as if they had come from unidentified remains. A facial approximation was generated using the known age, sex, and ancestry of the scan donor and the unadjusted, average weight as determined by