J Digit Imaging (2012) 25:347–351
DOI 10.1007/s10278-011-9429-3

Facial Recognition Software Success Rates for the Identification of 3D Surface Reconstructed Facial Images: Implications for Patient Privacy and Security

Jan C. Mazura · Krishna Juluru · Joseph J. Chen · Tara A. Morgan · Majnu John · Eliot L. Siegel

Published online: 8 November 2011

Abstract Image de-identification has focused on the removal of textual protected health information (PHI). Surface reconstructions of the face have the potential to reveal a subject's identity even when textual PHI is absent. This study assessed the ability of a computer application to match research subjects' 3D facial reconstructions with conventional photographs of their face. In a prospective study, 29 subjects underwent CT scans of the head and had frontal digital photographs of their face taken. Facial reconstructions of each CT dataset were generated on a 3D workstation. In phase 1, photographs of the 29 subjects undergoing CT scans were added to a digital directory and tested for recognition using facial recognition software. In phases 2–4, additional photographs were added in groups of 50 to increase the pool of possible matches and the test for recognition was repeated. As an internal control, photographs of all subjects were tested for recognition against an identical photograph. Of 3D reconstructions, 27.5% were matched correctly to corresponding photographs (95% upper CL, 40.1%). All study subject photographs were matched correctly to identical photographs (95% lower CL, 88.6%). Of 3D reconstructions, 96.6% were recognized simply as a face by the software (95% lower CL, 83.5%). Facial recognition software has the potential to recognize features on 3D CT surface reconstructions and match these with photographs, with implications for PHI.

Keywords Facial recognition · Privacy · 3D Reconstruction · 3D Imaging (imaging, three-dimensional) · HIPPA

J. C. Mazura (✉) · K. Juluru
Department of Radiology, Weill Cornell Medical College,
525 E 68th St, F-056,
New York, NY 10065, USA
e-mail: jmazura@gmail.com

M. John
Division of Biostatistics and Epidemiology, Department of Public Health, Weill Cornell Medical College,
402 E 67th Street,
New York, NY 10065, USA

J. J. Chen · T. A. Morgan · E. L. Siegel
Department of Diagnostic Radiology and Nuclear Medicine,
University of Maryland School of Medicine,
22 S Greene St,
Baltimore, MD 21201, USA

E. L. Siegel
Department of Diagnostic Radiology,
VA Maryland Healthcare System,
10 N. Greene Street,
Baltimore, MD 21201, USA

Introduction

Three-dimensional reconstructions of radiographic datasets are made possible by the use of advanced visualization tools that, with improvements in functionality and ease of use, have proliferated over the past decade. Three-dimensional surface rendering, when applied to datasets of the head and face, is able to demonstrate facial features (Fig. 1) [1]. It has been suggested that sufficient detail is included in these surface reconstructions to enable the identification of the person to whom the radiographic dataset belongs [2]. Although the potential to visually identify a patient is acceptable in situations of routine clinical care, it is generally not acceptable when radiographic datasets are used for research or publication or are made publicly accessible as educational materials.

The Privacy Rule, published in 2002 as a supplement to the Health Insurance Portability and Accountability Act (HIPAA) of 1996, mandates the security of protected health information (PHI), defined as "individually identifiable health information

Fig. 1 A digital photograph (*left*) and a 3D reconstructed image (*right*) of the same patient

that is transmitted or maintained in any form or medium by a covered entity or its business associates, excluding certain education or employment records." Individually identified information, according to this statute, includes "full-face photographic images and any comparable images" [3].

Medical data can be used in research settings if patient authorization is obtained. Alternatively, if PHI is removed and a randomized identification number is assigned, these medical data are considered de-identified and would not fall under HIPAA regulations. Even when textual information is removed from radiographic datasets, however, 3D surface reconstructions of the face may fall into the category of "full-face photographic" or comparable images. The problem is relevant not only for institutional imaging repositories but also for national repositories, such as the National Biomedical Imaging Archive. These can contain cross-sectional imaging files for thousands of patients and, although believed to be stripped of all identifying information, may contain the raw data required to create potentially identifiable 3D reconstructions of the face [4, 5].

It remains to be determined whether soft tissue reconstructions of the face have sufficient detail to identify the person to whom such an image belongs. At least one study by Chen et al. [2] tested the ability of human beings to match 3D reconstructed images with photographs of those same research subjects. The purpose of our study was to test the ability of a free and readily available computerized facial recognition application to match digital photographs of study participants with 3D facial reconstructions obtained from thin-section CT scans that included the face.

Materials and Methods

Study Overview

Data for the study, which was approved by the institutional review board, were acquired prospectively at an academic-affiliated Veterans Affairs (VA) Medical Center. Two groups of patients were recruited to form "reconstruction" and "control" arms of the study. The reconstruction arm (n=30) included outpatients undergoing scheduled CT scans of the maxillofacial sinuses or cerebral vasculature. After informed consent, the faces of all study participants were digitally photographed using a 6-megapixel camera (Lumix DMC-FX3K, Panasonic, Secaucus, NJ) with "MEGA" optical image stabilization. All digital photographs were acquired using a combination of image stabilization, flash, ×4 optical zoom, and portrait settings. The photographs were obtained at a distance of 4 ft using a standardized solid cream-colored background and an identical lighting source.

CT scans were performed according to standard clinical protocols with thin (0.75 mm) sections. CT datasets from the reconstruction arm were used to create 3D surface reconstructed images of each subject's face using an advanced visualization workstation (AquariusNET server, TeraRecon, Inc., San Mateo, CA). A standard color template with minimal window leveling was applied in order to create a standardized skin color.

The second set of study participants who comprised the control (photo-only) arm (n=150) was recruited from consecutive patients undergoing any type of scheduled outpatient radiologic imaging in the VA Maryland Healthcare System. After informed consent, all study participants in this arm were photographed using the same digital camera in the same manner as described previously. No 3D surface reconstructions were performed in this patient group.

Image Matching

A free and readily available computer image management application with facial recognition capabilities (Picasa 3.6, Google Inc., Mountain View, CA) was installed using a computer running the Mac OS X 10.5.8 operating system (Apple Inc., Cupertino, CA). The program automatically scans new images for faces and compares new images to existing faces previously imported. Thus, when facial images are imported, one of four events occurs: (1) the image is not recognized as a face (there is no prompt to name or confirm the identity of the image); (2) the image is recognized as a face with no matching image, and the user is prompted to name the “new” face; (3) a correct potential matching image is found, and the user is prompted to confirm the match; or (4) an incorrect potential matching image is found, and the user is prompted to confirm the match.

Image matching was conducted in four phases. In phase 1, frontal digital photographs of the subjects from the reconstruction arm were imported into Picasa and each given a unique numeric identifier. The 3D reconstructed image of each individual in the reconstruction arm was then imported sequentially, one at a time, and accuracy was recorded by noting which of the four recognition events listed previously occurred (Fig. 2). The 3D image was then deleted. In phases 2–4, additional photographs from the photo-only arm were imported into Picasa in groups of 50 (phase 2 had 50 additional images, phase 3 had 100, and phase 4 had 150), and the test for recognition was repeated for each phase as described previously.

As an internal control to test the most basic accuracy of facial recognition, photographs of all 180 subjects were tested for recognition against a duplicate uniquely named photograph in the same fashion outlined previously.

Statistical Analysis

Statistical analysis was performed using Microsoft Excel. Categorical variables, including ethnicity and sex, were compared between the control and study arms using a χ^2 test. Comparison of mean age between the two arms was performed using a t test, and 95% confidence limits were calculated using a normal approximation method.

Sample size considerations for the internal control group were based on obtaining a precise 95% confidence lower limit for the rate at which Picasa correctly identified a picture with its digital copy. Although the rate of correct detection in the internal control experiment was hypothesized to be 100%, we used a more conservative estimate of 90% for sample size considerations. For the internal control group, with a sample size of 30 and with the above assumption about the rate of correct detection by Picasa, the lower limit of a one-sided 95% confidence interval for this rate was projected to be 76%. The study was not powered for subgroup analysis based on sex or ethnicity.

Results

One of the 30 subjects recruited into the reconstruction arm was excluded because of corruption of the digital photograph. A total of 150 subjects were recruited into the photo-only arm.

Demographic data in the two arms demonstrated no statistically significant difference (Table 1). In the photo-only arm, the majority of study participants were African American (62%) men (80%), with a mean age of 53.5± 12.7 years. In the reconstruction arm, the distributions were similar, with a slightly lower percentage of African Americans (55%) and a slightly higher percentage of men (83%). The mean age in the “reconstruction” arm was 56.6± 13.1 years.

Of the 29 subjects in the reconstruction arm, all except one of the 3D reconstructions (96.6%) were recognized as a face. Eight of the 29 (27.64%) 3D images were matched correctly with the digital photographs of the subjects.

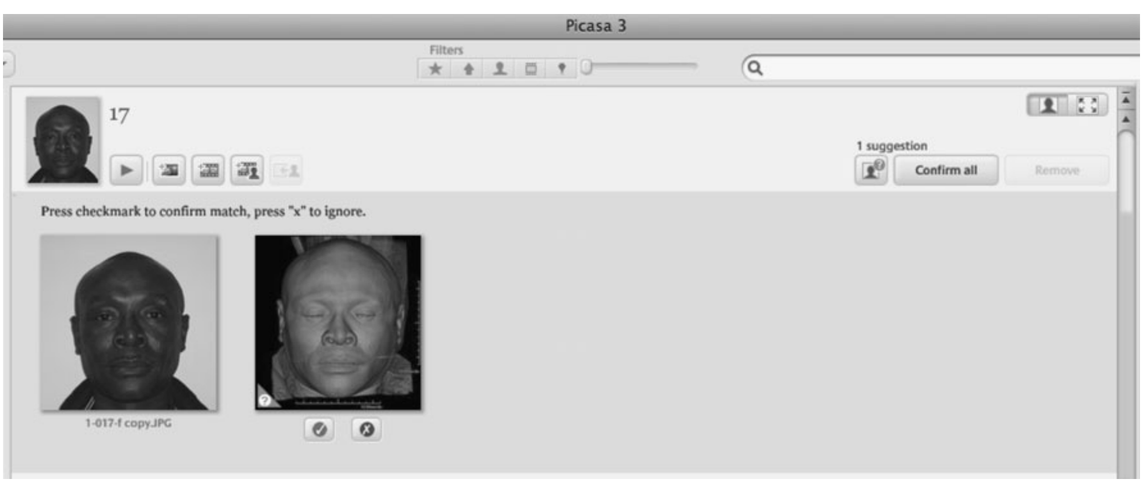

Fig. 2 Screen capture from Picasa in which Picasa correctly suggests that the person in the photograph and the 3D reconstructed image represent the same individual (event 3 as outlined in the text representing a correct match)

Table 1 Participant demographics

Demographic	Control arm (%) (n=150)	Study arm (%) (n=29)	P value
Ethnicity			0.802
African American	93 (62%)	16 (55%)	
White	55 (37%)	13 (45%)	
Asian	1 (1%)	0	
Pacific Islander	1 (1%)	0	
Sex			0.931
Men	120 (80%)	24 (83%)	
Women	30 (20%)	5 (17%)	
Age			0.269
Range (years)	21.8–88.5	37.3–84.8	
Mean+SD (years)	53.5±12.7	56.6±13.1	

Twenty-one of 29 (72.4%) 3D images were not matched with any digital photographs. None of the 3D reconstructions were matched with the wrong digital photograph. The matching pattern was identical across all phases, with reconstructions of the same subjects consistently matched correctly or consistently not matched at all. The overall accuracy of Picasa was 27.6%, which was independent of the number of photographs available for matching, up to a maximum of 179 (Table 2). In the internal control experiment, 100% of photographs of all subjects were matched correctly with a duplicate copy of that same photograph.

Discussion/Conclusion

These results demonstrate that soft tissue reconstructions of thin-slice head CTs have sufficient detail to enable computerized matching with subjects' photographs in slightly more than one out of four cases (Picasa software recognized 28/29 or 96.6% of the 3D reconstructed images as a face (95% lower CL) and correctly matched 27.6% of 3D images with the corresponding digital photograph). The success of the application was independent of the number of photographs available for matching, up to our maximum of 179 photographs. Although not every reconstruction could be matched with a photograph, the accuracy of Picasa was 100% when a match was proposed.

The overall accuracy rate of Picasa in our study was 27.6%, lower than the 57.2% accuracy rate of human reviewers reported in the Chen et al. report. However, the method of image matching in the Chen report was different, limiting a direct comparison of accuracy. Chen used a Web-based computer program that allowed image reviewers to match 3D reconstructed images of study participants with digital photographs of study participants in either the study or a control arm and also provided an answer option of "none of the above." Accurate matching occurred when a reviewer either correctly matched a 3D reconstructed facial image with its respective digital photographs or selected "none of the above" when the correct match was not available.

The future of facial recognition, especially when attempting to mine a large database of images, undoubtedly lies in facial recognition software. Using software for facial recognition is far less cumbersome and more efficient than having individual human reviewers. Moreover, we suspect that the accuracy of human reviewers would decrease and the time required for review would increase as the dataset comprising photos available for match is increased—in contrast to the performance of the computer algorithm.

Our findings suggest that subjects whose thin-slice head CT data reside in research repositories have the potential to be identified even when textual PHI is removed. Identifying individuals from their CT scan data would require the original CT data, soft tissue reconstruction tools, a collection of reference photographs, and facial recognition software.

The legal implications of this capability/technology should be further explored. Even with textual information removed from radiographic datasets, soft tissue reconstruc-

Table 2 Picasa facial recognition accuracy

	Phase 1	Phase 2	Phase 3	Phase 4
% reconstructions matched correctly (95% upper CL) to photographs	27.6 (40.1)	27.6 (40.1)	27.6 (40.1)	27.6 (40.1)
% reconstructions matched incorrectly	0	0	0	0

CL confidence limit

tions of the face may fall into the category of “comparable images,” and the potential of this technology calls into question whether these data can truly be considered de-identified. As facial recognition software improves, it will be increasingly important to develop tools or processes to protect the personal identity of patients who have undergone head imaging. Such technology may include software that distorts facial features or software that selectively removes surface soft tissue data from head CT datasets.

In 2006, Google acquired Neven Vision, a company specializing in facial recognition technology, and in 2008 incorporated Neven’s facial recognition software into Picasa. We choose Picasa not only because it was readily available and free but also because Neven Vision was among the top finishers in the 2002 and 2006 Facial Recognition Vendor’s Test, a government-sponsored independent competition comparing the world’s best facial recognition technologies [6].

Our study has limitations. First, this was a small study with a relatively small number of study participants in both the study and control arms. Second, the expressionless standardized photographs of the study participants used in this study cater to the strengths of facial recognition software. The human face is capable of an almost limitless number of expressions that distort facial features and hinder the accuracy of facial recognition software.

To our knowledge, this is the first study to evaluate computerized facial recognition of 3D reformatted radiology images. The study was performed in a controlled environment in which the subjects, the photography, and reconstructions were standardized, using a single facial recognition tool. Future research could focus on how recognition accuracy varies with photographic technique, ethnicity, sex, and using very large photo banks for comparison.

Acknowledgements We would like to thank Dr. Nancy Knight of the University of Maryland for help with editing the paper.

References

1. Marsh JL, Vannier MW: Three-dimensional surface imaging from CT scans for the study of craniofacial dysmorphology. J Craniofac Genet Dev Biol 9(1):61–67, 1989
2. Chen JJ, Khan SM, Fort L, Moffitt R, Juluru K, Kim W, Safdar N, Siegel EL: Observer success rates for identification of 3D surface reconstructed facial images and implications for patient privacy and security. SPIE Conference Proceedings Paper, 21 March 2007
3. Health Insurance Portability and Accountability Act. Available at: http://aspe.hhs.gov/admnsimp/pl104191.htm. Accessed 23 April 2011
4. Fetzer DT, West OC: The HIPAA privacy rule and protected health information: implications in research involving DICOM image databases. Acad Radiol 15(3):390–395, 2008
5. Avrin D: HIPAA privacy and DICOM anonymization for research. Acad Radiol 15(3):273, 2008
6. Phillips PJ, Scruggs WT, O’Toole AJ, Flynn PJ, Bowyer KW, Schott CL, Sharpe M: FRVT 2006 and ICE 2006 large-scale results. NISTIR 7408. National Institute of Standards and Technology, Gaithersburg, 2007