JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, January 2012, Vol. 57, No. 1
doi: 10.1111/j.1556-4029.2011.01944.x
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

ANTHROPOLOGY

Cristina Cattaneo,[1] M.D., Ph.D.; Angela Cantatore,[1,2] M.M.E.; Romina Ciaffi,[1] B.Sc.; Daniele Gibelli,[1] M.D.; Alfredo Cigada,[2] M.M.E.; Danilo De Angelis,[1] D.D.S.; and Remo Sala,[2] M.M.E.

Personal Identification by the Comparison of Facial Profiles: Testing the Reliability of a High-Resolution 3D–2D Comparison Model

ABSTRACT: Identification from video surveillance systems is frequently requested in forensic practice. The "3D–2D" comparison has proven to be reliable in assessing identification but still requires standardization; this study concerns the validation of the 3D–2D profile comparison. The 3D models of the faces of five individuals were compared with photographs from the same subjects as well as from another 45 individuals. The difference in area and distance between maxima (glabella, tip of nose, fore point of upper and lower lips, pogonion) and minima points (selion, subnasale, stomion, suprapogonion) were measured. The highest difference in area between the 3D model and the 2D image was between 43 and 133 mm^2 in the five matches, always greater than 157 mm^2 in mismatches; the mean distance between the points was greater than 1.96 mm in mismatches, <1.9 mm in five matches ($p < 0.05$). These results indicate that this difference in areas may point toward a manner of distinguishing "correct" from "incorrect" matches.

KEYWORDS: forensic science, personal identification, 2D/3D comparison, video surveillance system, laser scanner, facial profile

In the past few years, the request for expert opinions concerning identification of the living from pictures taken from video surveillance systems has increased. Facial identification of the living still, however, remains a difficult issue from a technical point of view. Indeed, the first attempts at identification of the living date back to the end of the 19th century with Alphonse Bertillon, who developed a morphological and metrical system for the description of facial features (1). From a practical and more modern point of view, there are three general approaches to facial identification: morphological classification of facial traits, metrical analysis, and the "face-on-face" superimposition methods. The first method is based on the classification of different facial features according to standardized classifications, which derive from the Interpol model (2–4). The morphological comparison of facial features should, however, be considered only a preliminary step, as it is subjective, with a large interobserver error (5). The metric assessment of faces is based on a quantitative analysis that is on the measurement of facial dimensions and indices; this approach however has been severely criticized by Kleinberg et al. (6) who verified the correspondence of four facial landmarks on an image taken from a video and a sample of 10 photographs, which included the photo of the subject represented in the video; all photographs were taken in the best conditions. Results indicated that the subject could not be accurately and reliably identified.

Thus, among various methods of facial assessment, the superimposition of images may turn out to be more reliable than morphological and metrical approaches; superimposition can be performed between 2D images of the culprit or criminal and the suspect (2D–2D comparison) or between the 2D image of the culprit taken from the video surveillance system, and a 3D virtual model of the suspect obtained by a 3D optical digitizer (2D–3D comparison). There are also 3D–3D comparison methods, but these require the use of 3D recording devices for video surveillance (such as 3D scanners). The 2D–2D comparison methods, however, require that the images represent the suspect in the exact same position as the culprit in the video surveillance image—and this may be difficult to achieve (7–10). Moreover, 2D–2D procedures show some intrinsic limitations because recognition efficiency mainly depends on different variables (such as illumination, face positioning and orientation).

The 2D–3D comparison method, therefore, seems to be more reliable for personal identification compared with other morphological and metrical approaches and brings on several advantages, such as the recording of 3D facial models and the creation of specific image databases for one-to-many comparison procedures (7). The acquisition is obtained by a laser scanner, and the reconstruction of the 3D model achieved by merging acquired images; the comparison with 2D images is performed by resizing, repositioning, and reorienting single images with specific software, until the best match with the culprit's face is achieved (8).

The 3D model of the suspect and superimposition on the 2D image of the culprit allows examiners to obtain an easy-to-read comparison and to assess how closely the profiles match (8). However, a

[1]LABANOF, Laboratorio di Antropologia e Odontologia Forense, DMU—Dipartimento di Morfologia e Scienze Biomediche, Sezione di Medicina Legale, Università degli Studi di Milano, Milan, Italy.

[2]Sezione di Misure e Tecniche Sperimentali, Dipartimento di Meccanica, Politecnico di Milano, Milan, Italy.

Received 13 June 2010; and in revised form 7 Oct. 2010; accepted 29 Oct. 2010.

standardized method for assessing identification from the 3D–2D comparison has not yet been developed to correctly distinguish matches from nonmatches, nor has an attempt at quantifying differences been made.

The issue of "quantification" in identification is a questionable one: some researchers conclude that it is useless; others, on the contrary, search for some sort of confort in numbers. Regardless of whether one agrees or disagrees with the need to "quantify" a result, we cannot ignore the fact that frequently judges and magistrates expect a result expressed as frequencies and/or probabilities. If we consider a face, its uniqueness seems intuitive; however, it is difficult (and may be impossible) to express this from a quantitative point of view. On the other hand, in the forensic scenario, every method requires a known error, and in this particular case, the quantified probability that two profiles, or faces in general, belong to the same individual could be an important judicial tool. Thus, it seems reasonable to verify whether this is possible at all. At the moment, there is no quantitative indication concerning how stringent a match between two profiles from the same individual may be, or how different those from different individuals are, although recent research has underlined this problem, and in particular how even the influence of sex, ancestry, age, and other variables must be taken into account (11–13).

There are some attempts in the literature to standardize 3D–2D comparison techniques; the first attempt at quantifying the match between the 3D and 2D images was made by Yoshino et al., who measured the distance between correspondent facial points taken on a 3D model and a photo taken from the same individual. The authors found that the mean distance of 14 facial points between the two images was 1.3 mm, the highest error amounting to 2.9 mm, the lowest to 0.0 mm (7). Yoshino et al. tried this first method of quantifying matches between 3D and 2D images in cases of 2D images reproducing disguised individuals (9). 3D facial data from 100 Japanese males were compared with photos reproducing three subjects wearing a pair of sunglasses, a cap, and a gauze mask; the matching criterion was the sum of the average point-to-point difference of the corresponding landmarks between the 3D and 2D images. When the 3D model and 2D image were from the same individual, the sum of point-to-point differences was 1.43–3.3 mm, with an overall difference between the two outlines amounting to 0.5–1.6 mm; in case of nonmatches, the point-to-point difference was 2.6–7.0 mm, with a difference between the two outlines ranging between 2.6 and 7.0 mm. The authors pointed out the high reliability of the morphometrical matching method for personal identification and the possibility of distinguishing correct matches from incorrect ones (9).

A different point of view was expressed by Goos et al., who tried to quantify correspondence between 3D models and 2D images by measuring distances between facial points (10). In detail, a set made of a 3D model and a photo from the same person as well as a 3D model from a second individual was created; seven facial points were chosen and correspondence between the 3D model and images was tested, in the attempt to match all seven points (the error was measured as distance between the different points in 3D and 2D images) or only four points (the error was measured as distance of the three corresponding points not employed in the match). Results showed that the distance between the matching points is usually higher in matching pairs than in nonmatching pairs of images, whether four or seven points are used. The authors conclude that anthropometric comparison of 3D–2D images cannot be used as an identification method, although they admit that more research needs to be carried out with a wider sample of facial comparisons.

As one can observe, the studies concerning face-on-face superimposition are trying to find a common quantitative way for expressing the concordance rate between the 3D and 2D images. However, very few studies are available in literature, and there is no common manner of analyzing superimposition. Thus, the present study aims at devising a method of quantifying the correspondence between 3D and 2D images based on the comparison of a simple projection of the face, the profile, derived from 2D images and 3D models, and calculating the mean differences between two facial profiles obtained from the same person; in other words, this brief pilot study strives to verify whether one can find an answer to the question "how (if possible) can an expert in court quantify the probability that two different persons share the same profile?"

Materials and Methods

A group composed of 50 young Caucasian men (aged between 20 and 30 years) was selected, with no mimicking facial characters such as beards. A high-definition digital camera was used to control all environmental conditions and validate the method in nearly ideal conditions. Frontal, right and left profile pictures were taken from each face belonging to the database. The photo of the right profile of each subject was chosen for the study.

Five subjects within the database were further selected to obtain a 3D model of their face via a laser scanner (Konica Minolta model VI-9i, Tokyo, Japan), assuring an accuracy of ±0.05 mm. The device had already been employed in a previous study (8) in which the authors standardized the procedure of 3D–2D superimposition with the same equipment. From each 3D model, the corresponding profile was defined to perform the comparison with the 2D image.

The profile (traced between the trichion and gnathion) of each subject belonging to the photographic database was then compared with the right profile of each 3D model of the five scanned subjects. Each profile from the 3D model was superimposed to the profiles of the 50 photographs within the database; 250 comparisons were, thus, performed, among which 245 were between 3D models and photographs of different subjects, named "incorrect matches," and five between 3D models and photographs of the same individuals, named "correct matches." Tests were all classified as blind because the operator who performed the superimposition did not know which were the true correct matches. The comparison was obtained by projecting the 3D facial model of each scanned subject on the 2D picture of each subject of the database, reaching the best match by following guidelines reported in a previous study by the same authors (8) (Figs. 1 and 2). (Briefly, in this previous study [8], the authors apply the concept of geometrically compatible images. A 3D facial model is compared to a photograph by superimposition. Repositioning and reorientation of the 3D model according to the photograph are manually accomplished after automatic resizing.) Then, matching was assessed by a quantitative analysis starting from the superimposed profiles; differences in the areas described by the two profile lines (the profile from the 3D model and the photo of the right profile) and differences between maxima and minima points extracted from the two profile curves were evaluated. The maxima points chosen were glabella, tip of the nose, the fore point of upper and lower lips, and pognonion; the minima points chosen were selion, subnasale, stomion, and suprapogonion.

Profile curves, in terms of contour image coordinates (x, y), were computed by elaborating the original 3D picture and the image of the 3D model projected. Profile curves were automatically extracted by implementing a program in Microsoft Visual Basic 6.0,

FIG. 1—*Example of correct matching between the 3D model and photograh profiles.*

FIG. 2—*Example of incorrect matching between the 3D model and photograph profiles.*

provided with Matrox Imaging Libraries (MIL) for image processing. Edge detection was applied. This image processing operator provides contour point extraction by analyzing gray values of each pixel. It works on a one-image channel (8-bit images) and, in case of color images, it requires either to convert them to gray or to consider a one-color channel (red, green, or blue one). To maximize technique reliability, it is generally advisable to binarize the image, in other words to convert each pixel value to only two values, white or black, thus highlighting the target object.

To extract profile coordinates of the 2D image, it is useful to process the red color channel, because skin, usually pink in this case, has higher red color values if a proper background is chosen (in this case, a blue background was applied). Concerning the 3D model projection image, the model in green was chosen and a black background was adopted to apply the edge detection analysis to the green color channel image.

In Fig. 3, profile curves refer to a "match" case where the 3D model and 2D image derived from the same individual, and a "mismatch" case where two profiles belonging to different subjects are shown.

The concordance within each pair of profiles was evaluated by computing the difference between the areas described by the two curves and the difference between the maxima and minima points (Fig. 4). In all 250 comparisons, the difference in area and distance between the five maxima and four minima points of the two profiles were measured. Data from the 245 incorrect matches were compared with results from the five correct matches and underwent statistical analysis.

Results

In the 245 mismatches (where the photograph and the 3D model did not derive from the same subject), the mean difference in areas between the profile from the 3D model and the photo of the right profile was 1731.1 mm^2, whereas in the five correct matches (of the five 3D models and the photographs of the same individuals)

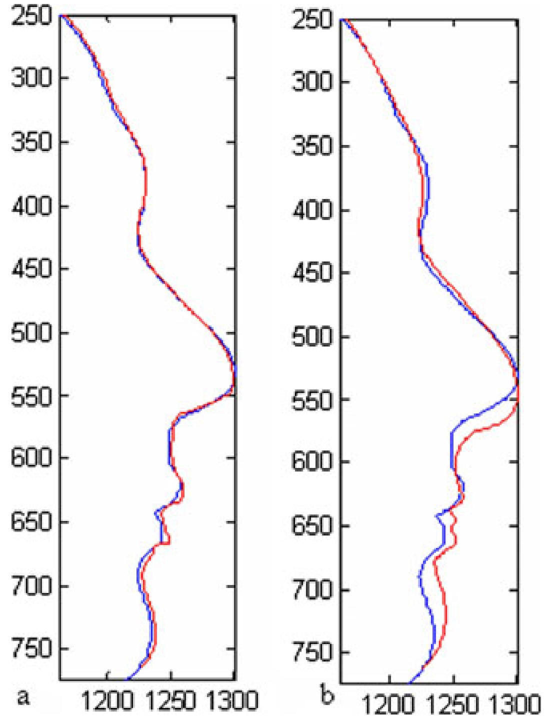

FIG. 3—*Profile curve in a "match" case (a) and "nonmatch" case (b). The profile curve extracted from the picture is represented in blue, while that extracted by the model projection is represented in red.*

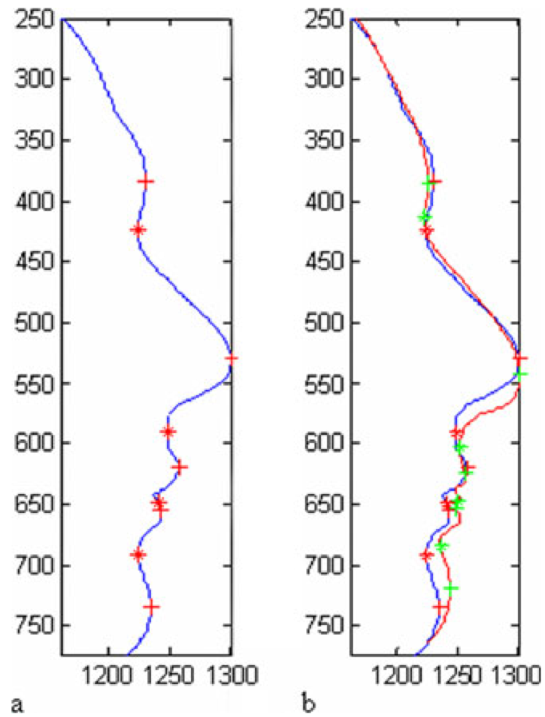

FIG. 4—*Comparison between minima and maxima points in 3D model and photograph profiles: in a "match" case (a) and "nonmatch" case (b).*

the difference was significantly lower, amounting to 87.4 mm^2 ($p < 0.05$). The highest difference in area was between 43 and 133 mm^2 in the correct matches, whereas in incorrect matches (where the photo and 3D model belonged to different subjects) the difference in area between the two profiles was always greater than 157 mm^2. In 8% of cases, the difference in area was between 157 and 299 mm^2, in 44% between 300 and 999 mm^2, in 48% the difference was between 1000 and 9100 mm^2 (Fig. 5).

In the 245 cases of mismatch (incorrect matches), the mean distance between the maxima points of the 3D model and 2D image was always higher than 1.73 mm, with a mean of 5.33 mm, whereas in the five correct matches the difference varied from 0.7 to 1.9 mm, with a mean of 1.28 mm ($p < 0.05$). In the cases of 245 mismatches, only in 1.6% the difference was between 1.73 and 1.9 mm; in 48.6% the difference was between 2.00 and 4.9 mm and in 49.8% between 5.00 and 15.5 mm (Fig. 6).

In the 245 mismatches, the mean distance of the minima points between the 3D model and 2D image was higher than 1.69 mm, with a mean of 6.49 mm; in the five correct matches the difference varied from 0.6 and 1.8 mm, with a mean of 1.20 mm ($p < 0.05$). For the 245 mismatches, only in 6.5% of cases the difference was between 1.69 and 2.99 mm; in 84.5% the difference of minima points was between 3.00 and 10.9 mm and in 9% between 11.00 and 18.00 mm (Fig. 7).

If one considers all points used for the comparison (both the maxima and minima ones), the difference between the 3D model and the photo from the same individual was lower than 1.9 mm, with a mean of 1.24 mm. In the 245 mismatches, differences between the two profiles were always greater than 1.96 mm, with a mean of 5.85 mm. Only in 5.7% of cases, the difference was between 1.96 and 2.99 mm; in 87.7% it was between 3.00 and 9.99 mm and in 6.6% higher than 10.00 mm (Fig. 8).

Discussion

This study aimed at quantifying the degree of profile matches in cases of personal identification of the living through 3D–2D superimposition methods. The face-on-face comparison seems to have surpassed the morphological and metrical methods for reliability and user-friendliness; however, the lack of a precise indication of error and probability rate for personal identification is a relevant limit, which has been faced by literature only in the last years (7,9,10,12).

This study presents a method of validation of 3D–2D superimposition techniques by quantifying the difference between areas and distances of maxima and minima points of facial profiles from a 3D model and a 2D image. The statistical analysis showed that the difference in areas seems to be a reliable piece of information for evaluating correct identification: in noncorrect matches the difference in area was always >157 mm^2, whereas in correct matches the highest difference was between 43 and 133 mm^2. A similar result was obtained by the overall difference of maxima and minima points between the 3D model and the 2D image: in cases of nonmatches the difference was always higher than 1.96 mm, whereas in the five correct matches this value was always lower than 1.9 mm. For both areas and distances of all points, the incorrect and correct matches showed no superimposition of values and therefore seemed to provide a valid starting point for the diagnosis of identification. The comparison of areas and point-to-point differences seems to be the most reliable indicator concerning correct identification; in addition, the different results highlighted by the analysis of maxima and minima points seem to support the hypothesis that the number of facial points compared is crucial for determining the significance of correct identification. In fact, the chosen minima points were only four, and the differences between the 3D model and 2D images in

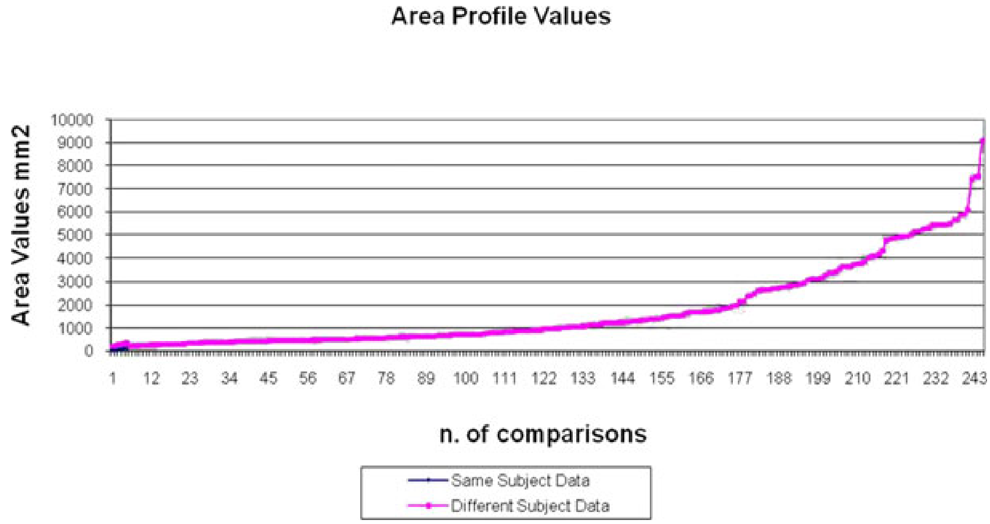

FIG. 5—*Difference in area between the photographs and 3D model in the 245 cases of mismatches.*

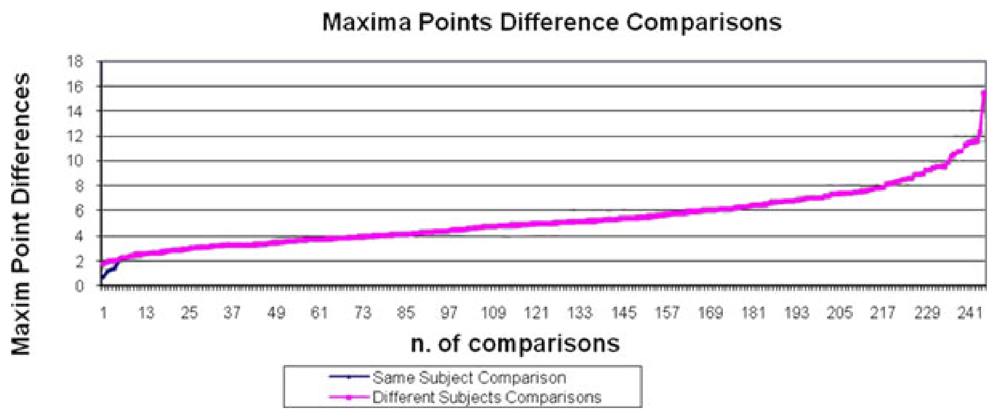

FIG. 6—*Distance of the maxima points between the 3D model and 2D image in the 250 comparisons.*

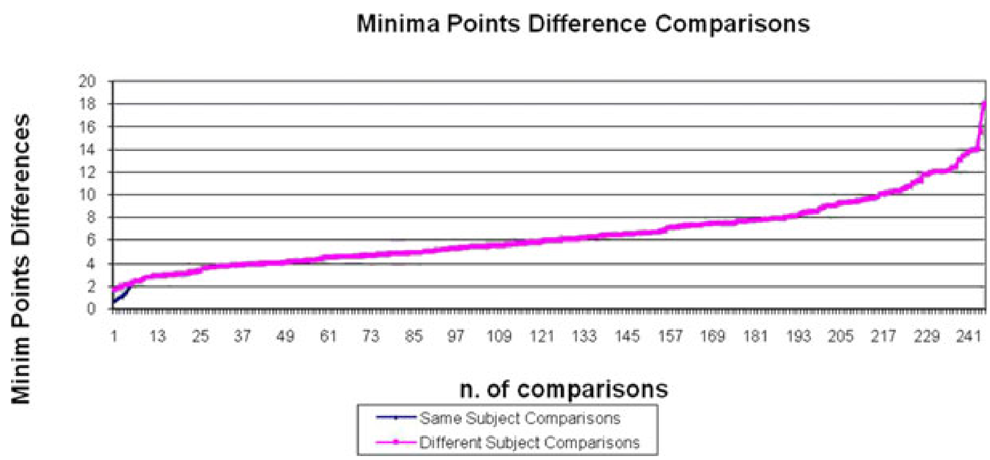

FIG. 7—*Distance of the minima points between the 3D model and 2D image in the 250 comparisons.*

cases of nonmatches and matches showed a range of superimposition in value: in correct matches the difference was between 0.6 and 1.8 mm; nonmatches showed a 1–2.99 mm difference in 6.5% of cases. On the other hand, the five maxima points showed a higher significance in difference between the 3D model and the 2D photograph: in correct matches the difference was between 0.7 and 1.9 mm, whereas in noncorrect matches the difference was less than 1.9 mm only in 1.6% of cases. This seems to prove that the increase in number of points of comparison raises the significance of differences between the 3D and 2D profiles and therefore the probability of a distinction between correct and incorrect matches. This may probably explain the different results

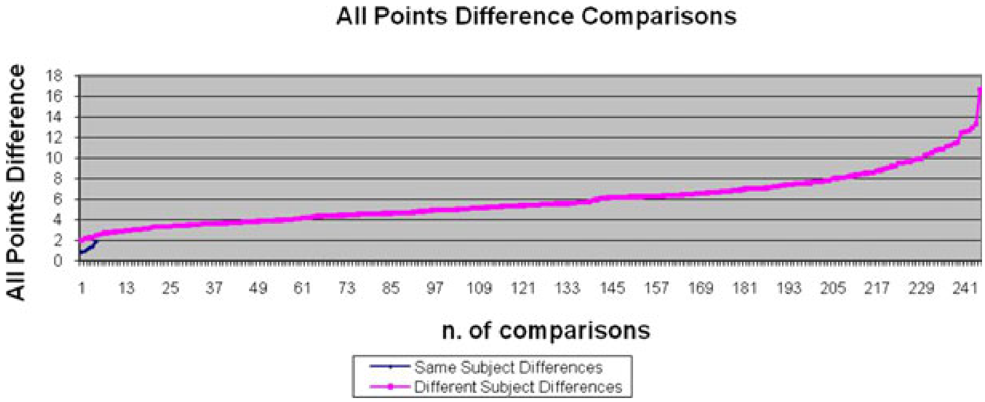

FIG. 8—*Distance of maxima and minima points between the 3D model and 2D image in the 250 comparisons.*

obtained by Goos et al. (10) and Yoshino et al. (7,9): the lower number of facial landmarks used for comparison by Goos et al. may have reduced the significance of point-to-point differences. Therefore, the negative results obtained by these authors may be explained by the low number of chosen landmarks; this suggests that a higher number of landmarks might increase the chances of a more precise distinction between matches and mismatches. Yoshino et al. in fact compared 14 points and were able to increase the distinction between correct and incorrect matches. In addition, intrinsic difficulties of placing landmarks on a 3D image may contribute to the increasing discrepancies between studies.

The results obtained by the present study however are not comparable with previous ones because our study aims toward the standardization of a 3D–2D comparison in a simplified model, such as that of a profile, whereas both Goos et al. and Yoshino et al. analyzed images in different orientations and measured different variables. From this point of view, the present study represents a simplified and ideal scenario (the profile). Nonetheless, the results are consistent with those reported by Yoshino et al., which may suggest that quantification of a profile match may be possible and reliable in different orientations of the face. Further studies are, however, necessary. At the present, orientation of the head in such images is still a crucial issue, and a method of validation of personal identification needs to be applicable in every situation, by selecting the more appropriate and reliable facial landmarks or profiles for a specific position of the culprit's head in the 2D images.

This study aimed simply at trying to verify how to quantify a profile match in the perspective of individual identification and at trying to assess the significance of a mismatch (or match), and the possibility that two different individuals share the same profile. This seems to be possible, though more research needs to be performed, particularly with respect to different orientations where outlines and profiles of the face are more difficult to compare.

In conclusion, this study has provided some insight into the quantification of the similarity/differences of profiles, though tests on larger samples are necessary.

Acknowledgments

The authors declare that the experimental project complied with the current laws of the countries in which it was performed.

References

1. Bertillon A. Signaletic instructions including the theory and practice of anthropometrical identification. Chicago, IL: National Library of Medicine, 1896.
2. Vanezis P, Lu D, Cockburn J, Gonzales A, McCombe G, Trujillo O, et al. Morphological classification of facial features in adult Caucasian males based on an assessment of photographs of 50 subjects. J Forensic Sci 1996;41(5):786–91.
3. Ohlrogge S, Nohrden D, Schmitt R, Drabik A, Gabriel P, Ritz-Timme S. Anthropological atlas of male facial features. Frankfurt, Germany: Verlag für Polizeiwissenschaft, 2008.
4. Ohlrogge S, Arent T, Huckenbeck W, Gabriel P, Ritz-Timme S. Anthropological atlas of female facial features. Frankfurt, Germany: Verlag für Polizeiwissenschaft, 2009.
5. Ritz-Timme S, Gabriel P, Obertovà Z, Boguslawski M, Mayer F, Drabik A, et al. A new atlas for the evaluation of facial features: advantages, limits and applicability. Int J Legal Med 2011;125(2):301–6.
6. Kleinberg KF, Vanezis P, Burton AM. Failure of anthropometry as a facial identification technique using high-quality photographs. J Forensic Sci 2007;52(4):779–83.
7. Yoshino M, Matsuda H, Kubota S, Imaizumi K, Miyasaka S. Computer-assisted facial image identification system using a 3-D physiognomic range finder. Forensic Sci Int 2000;109:225–37.
8. De Angelis D, Sala R, Cantatore A, Grandi M, Cattaneo C. A new computer-assisted technique to aid personal identification. Int J Legal Med 2009;123:351–6.
9. Yoshino M, Noguchi K, Atsuchi M, Kubota S, Imaizumi K, Thomas CDL, et al. Individual identification of disguised faces by morphometrical matching. Forensic Sci Int 2002;127:97–103.
10. Goos MIM, Alberink IB, Ruifrok ACC. 2D/3D image (facial) comparison using camera matching. Forensic Sci Int 2006;163:10–7.
11. Evison M, Dryden I, Fieller N, Mallett X, Morecroft L, Schofield D, et al. Key parameters of face shape variation in 3D in a large sample. J Forensic Sci 2010;55(1):159–62.
12. Mallett X, Dryden I, Vorder Bruegge R, Evison M. An exploration of sample representativeness in anthropometric facial comparison. J Forensic Sci 2010;55(4):1025–31.
13. Evison M, Vorder Bruegge R. Computer-aided forensic facial comparison. London, UK: Taylor & Francis Ltd., 2010.

Additional information and reprint requests:
Cristina Cattaneo, M.D., Ph.D.
Professor
LABANOF, Laboratorio di Antropologia e Odontologia Forense
DMU—Dipartimento di Morfologia e Scienze Biomediche
Sezione di Medicina Legale
Università degli Studi di Milano
V. Mangiagalli 37
Milan
Italy
E-mail: cristina.cattaneo@unimi.it