

TECHNICAL NOTE**ANTHROPOLOGY**

Marta R. P. Flores,¹ M.Sc.; Carlos E. P. Machado,² Ph.D.; Matteo D. Gallidabino ,³ Ph.D.; Gustavo H. M. de Arruda,⁴ Ph.D.; Ricardo H. A. da Silva,⁵ Ph.D.; Flávio B. de Vidal,⁶ Ph.D.; and Rodolfo F. H. Melani,¹ Ph.D.

Comparative Assessment of a Novel Photo-Anthropometric Landmark-Positioning Approach for the Analysis of Facial Structures on Two-Dimensional Images*

ABSTRACT: Positioning landmarks in facial photo-anthropometry (FPA) applications remains today a highly variable procedure, as traditional cephalometric definitions are used as guidelines. Herein, a novel landmark-positioning approach, specifically adapted for FPA applications, is introduced and, in particular, assessed against the conventional cephalometric definitions for the analysis of 16 landmarks on ten frontal images by two groups of examiners (with and without professional knowledge of anatomy). Results showed that positioning reproducibility was significantly better using the novel method. Indeed, in contrast to the classic approach, very low landmark dispersions were observed for both groups of examiners, which were usually below the strictest clinical standards (i.e., 0.575 mm). Furthermore, the comparison between the two groups of examiners highlighted higher dispersion consistencies, which supported a higher robustness. Thus, the use of an adapted landmark-positioning approach proved to be highly advantageous in FPA analysis and future work in this field should consider adopting similar methodologies.

KEYWORDS: forensic science, facial analysis, anthropometry, cephalometry, facial identification, facial image

Facial photo-anthropometry (FPA) is the sub-field of physical anthropology that deals with the systematic study and measurement of human facial traits from two-dimensional images (1–3).

¹Social Department of the Dentistry College, University of Sao Paulo, Avenida Professor Lineu Prestes 2227, Sao Paulo 05508-000, Brazil.

²Brazilian Federal Police, National Institute of Criminalistics, INC Building, Brasília 70610-200, Brazil.

³Centre for Forensic Science, Department of Applied Sciences, Faculty of Health and Life Sciences, Northumbria University, Ellison Building, Newcastle Upon Tyne, NE1 8ST, U.K..

⁴Brazilian Federal Police, Technical and Scientific Sector, Rua Annita Luiza Mello Di Lascio, Cabedelo 58101-770, Brazil.

⁵Department of Stomatology, Public Health and Forensic Odontology, Dentistry College of Ribeirao Preto, University of Sao Paulo, Avenida do Café, Ribeirao Preto 14040-904, Brazil.

⁶Department of Computer Science, University of Brasilia, Darcy Ribeiro Campus, Asa Norte, Brasilia 70910-900, Brazil.

Corresponding author: Marta R. P. Flores, M.Sc. E-mail: mrpflores@gmail.com

*Presented at the 21st Triennial Meeting of the International Association of Forensic Sciences (IAFS), August 21 - 25, 2017, in Toronto, Ontario, Canada; and at the 8th European Academy of Forensic Science (EAFS), August 27–31, 2018 in Lyon, France.

Funded and supported by the Ministry of Education, Higher Education Personnel Improvement Coordination (CAPES), Brazil (grant number: 25/2014 – Pro-Forenses).

This research was approved by the Research Ethics Committee of the Medical School of Ribeirão Preto, University of São Paulo (protocol number 17902713.7.0000.5419).

Received 6 Aug. 2018; and in revised form 12 Sept. 2018, 2 Oct. 2018; accepted 3 Oct. 2018.

Since facial measurements have been correlated with several individual characteristics, FPA has found large applications in a number of scientific fields in which the analysis of faces on two-dimensional images is of interest (1,3,4). In legal medicine and forensic science, in particular, different studies reported the possibility of using FPA to estimate the age of individuals (5–7), to predict their sex or ancestry (8,9), to simulate facial growth or age progression (7,10), as well as to support human identification by comparing captured facial images to reference ones, that is, forensic facial identification (FFI) (1,11,12).

The first step in every FPA application involves the placement of a number of reference points (i.e., landmarks) on the facial images of the analyzed individuals, which is a process conventionally performed by following definitions used in classic facial anthropometry (or, as it is also called, cephalometry) (1,3,13). Traditional cephalometric definitions, however, merely describe a series of purely anatomical structures lying on the skin surface and/or the underlying bones and were primarily established for the purpose of directly mapping actual living subjects or their lateral-view X-ray image for medical purposes (14–16). Consequently, their adoption in FPA applications usually leads to a high positioning variability within and between examiners (17–22). The main reason for this arises from the fact that different examiners may have different interpretations of where a specific cephalometric landmark should be placed on a two-dimensional, frontal view, facial image, without any three-dimensional reference and/or the possibility to touch the subject's actual facial surface. As a result of this, the general reliability of FPA has been recently challenged by the

scientific community (17–19,23). One significant aftermath, in particular, has been the recommendation from the Facial Identification Scientific Working Group (FISWG) to avoid using FPA-based methodologies as proof of evidence in FFI (21).

Even if it is acknowledged that the application of FPA-based methodologies may be difficult and inadequate in a number of situations, such as those involving low resolution and/or non-frontal facial images, in several others it is not and may actually be beneficial. This is the case, for example, in those situations where images are acquired under sufficiently standardized conditions, such as in the detection of identity document fraud or age estimation from portrait images (17,20,23,24). To guarantee highly reliable results, however, a high reproducibility in landmark location is still essential and improvements would therefore be necessary (22,23,25). In particular, it is advised that the aforementioned reproducibility issues may be reduced through the use of proper landmark descriptions and/or locating procedures optimized for FPA applications, which thus take into account the specific problems encountered when positioning landmarks on two-dimensional facial images.

Despite the numerous works in FPA, however, none have previously proposed this kind of adapted protocol, leaving a gap in the specialized literature. Recently, a novel FPA-specific landmark approach was suggested by Flores et al. (28). In addition to a complete series of descriptions for landmarks based on visual references, the work also included optimized operational procedures and illustrations to locate each landmark of interest on two-dimensional images. These are intended to better assist examiners in FPA analysis and thus improve both the reproducibility and robustness of the landmark placement procedure. The approach has nonetheless never been assessed. Consequently, the current work aimed to undertake this and, in particular, to evaluate the improvement in reliability from using this adapted approach (hereafter, AdMet) over the classic, cephalometry-based one (hereafter, CIMet).

In order to achieve these aims, the two approaches were applied to a set of ten frontal view facial images and variability of the placement of specific landmarks between different examiners (i.e., reproducibility) investigated through their spatial dispersions around the grand means. Two groups of examiners, composed of individuals with and without specific knowledge of anatomy, respectively, took part in the experiment. This was done in order to assess the robustness of the approaches with respect to the experience level of the examiner. Observed landmark dispersions were finally compared to clinical standards currently accepted in cephalometry, by converting pixel-based values to millimeters through iris ratio calibration (7,26,27). To our knowledge, this is the first time that the adapted, FPA-optimized landmark-positioning previously reported by Flores et al. (28) has been evaluated in published literature. It is also the first time that a comparative study between different landmark-positioning approaches for FPA analysis has been carried out, as well as that their relative reliabilities have been investigated and validated against previously reported clinical standards.

Materials and Methods

Reference Facial Images

Ten frontal view facial images (from five male and five female subjects) were randomly selected from a larger database composed of 500 Brazilian frontal view images. For capture, subjects were asked to adopt a neutral facial expression and their faces were aligned with the Frankfurt plane. All the two-dimensional

images were acquired using a Geometrix FaceVision® FV802 Series Biometric Camera (ALIVE Tech, Cumming, GA), with no interchangeable lenses, and positioned at 1.2 m from the individual's face, at a resolution of 1200 × 1600 pixels.

FPA Analysis

Two groups of examiners were selected. The first group, named experts group (EG), was composed of five examiners with specific knowledge of anatomy (master or doctoral students in medical or dental areas), as well as previous experience in anthropometry and/or cephalometry. The second group, named non-experts group (NG), was composed of five examiners with higher education in scientific fields out of medical sciences, with neither training or specific knowledge of anatomy nor previous experience in anthropometry and/or cephalometry.

Both groups were asked to map the previously selected facial images according to two different landmark-positioning approaches: a classic method (CIMet) and a newly developed adapted method (AdMet). Generally, the mapping involved placing 16 specific landmarks on facial images, 8 odd (medians), and 8 even (laterals), as shown in Fig. 1. For CIMet, examiners were provided with a list of definitions for the 16 landmarks, previously compiled from a set of particularly influential works in craniofacial anthropometry (29–31) (Table 1). For AdMet, examiners were provided with the respective definitions and operational marking procedures obtained from the work of Flores et al. (28). This approach has been translated into a manual that is publicly available at http://facisgroup.org/facial_landmarks and included in Appendix S1.

The AdMet approach provides the examiner with clearer reference points that explicitly mention visible facial features instead of being solely based on anatomical structures. Furthermore, each described facial landmark includes a brief operational procedure and graphical illustrations, intended to better support locating it on images. The difference between CIMet and AdMet can easily be highlighted through an example. The ectocanthion landmark is conventionally defined as: “*the lateral corner (angle) of the eye*” (29–31). The newly adapted approach (28), on the contrary, reports the following definition: “*The most lateral landmark in the corner of the eye (distant from the midline), where the upper and lower ciliary implantation lines meet*” (p. 07). The following positioning procedure is also provided: “*Move the vertical line from lateral to medial side of the face to the landmark where the upper and lower ciliary lines meet in the region of lateral angle of the eye. Then, move the horizontal line until the point of convergence of those lines. Mark ectocanthion in the intersection region between the two auxiliary lines*” (p. 07). See Fig. 2 for an illustration of the corresponding page of the manual (SI). The manual describes a total of 36 landmarks. For the sake of comparison, however, only the 16 for which cephalometric definitions could be applied were selected in this work.

The FPA analyses with the two different landmark-positioning approaches were carried out by the same participants, with a month interval in between (starting from CIMet), in order to minimize memory effects on landmark placement. For each approach, examiners were asked to analyze the same 10 facial images in triplicate, again with a week interval in between. For mapping, a non-commercial software package for two-dimensional facial analysis was used, that is, SAFF-2D® (Forensic Facial Analysis System, Department of Federal Police, Brazil). The software allows examiners to locate the facial landmarks on images and to automatically register them through Cartesian coordinates (X, Y).

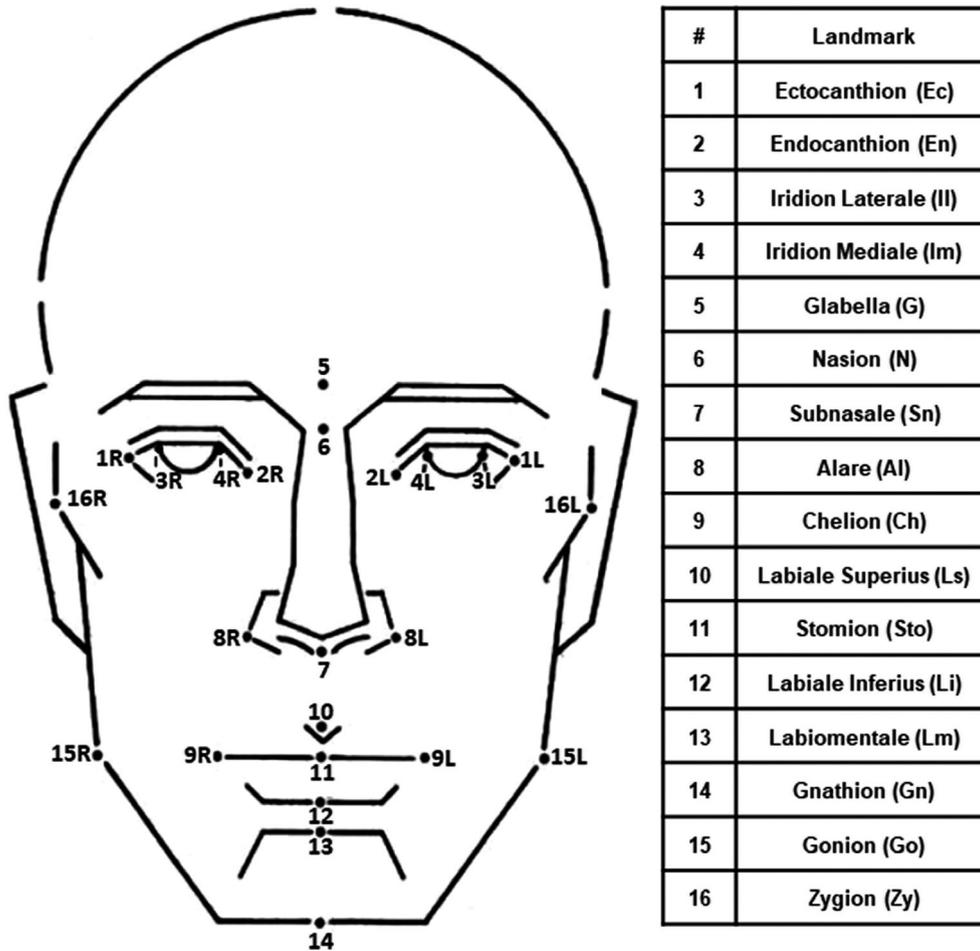


FIG. 1—Facial diagram representing the 16 landmarks used (left) and their nomenclature (right). Letter R corresponds to the right side and L to the left side.

Data Treatment

Initially, for each replicate experiment, average coordinates for all 16 landmarks were calculated for the three analyses. Then, differences (in pixels, px) on both the horizontal and vertical axes were determined between these average coordinates and the grand between-faces means. Location dispersions were defined as the mean differences on the horizontal axis (D_X) and mean differences on the vertical axis (D_Y). The arithmetic mean between these two values, that is, the mean dispersion (D_{MXY}), was also determined as summary statistics (17).

Values for D_X , D_Y , and D_{MXY} were then converted into an actual physical scale (i.e., from px to mm) by applying a scaling factor of 4.35 px/mm, in order to allow comparison of observed dispersions with previously published clinical standards. This scaling factor was previously determined by size comparison of a reference anatomical structure measured from images and real persons. The iris diameter was used for this purpose, as it has previously been proved to be an adequate reference for facial image calibration (7,26,27). Considering that the average iris diameter in images was calculated to be around 50 pixels and that the maximum population value of the horizontal visible iris diameter (HVID) is described in specific literature as around 11.5 mm (32–34), a ratio of 4.35 px/mm was determined. Converted dispersions were referred to as “estimated real dispersions” (ERD), that is, ERD_X , ERD_Y , and ERD_{MXY} (17).

Results Assessment

The normality of the data was initially assessed by the Shapiro-Wilk test and the intra-examiner marking reliability by the intra-class correlation coefficient (ICC). Analysis of variance was applied to assess any significant differences between the dispersion values resulting from the tested factors (i.e., the expert groups and FPA protocols). This was performed using marginal linear regressions with a gamma distribution for the errors. Results of all these statistical analyses were assessed against a statistical significance level of 5% ($\alpha = 0.05$).

For clinical validation, ERD values were compared against reference thresholds previously reported in the literature. In this respect, values smaller than 0.575 mm were considered ideal, based on the most strict references in cephalometry (13) (mean between 0.59 mm and 0.56 mm), while values between 0.575 and 1 mm were considered acceptable (14,25,35–38). ERD values >1 mm were considered undesirable.

Results

General Statistical Analysis

Firstly, the normality of the data was assessed. The Shapiro-Wilk test indicated that the data were not normally distributed and, thus, a non-parametric statistical analysis was subsequently conducted. The ICC test results showed that the

TABLE 1—List of the 16 investigated facial landmarks, with the corresponding sets of adopted cephalometric and facial photo-anthropometric (FPA) descriptions (used in the CIMet and AdMet landmark-positioning approaches, respectively). Cephalometric descriptions were compiled from those reported by George (29), Kolar and Salter (30), and Zimble and Ham (31). FPA-specific descriptions were extracted from the FPA manual provided in the Appendix S1 and the corresponding pages are reported in the table

#	Landmark	Abbr.	Cephalometric Description (CIMet)	FPA Description (AdMet)
1	Ectocanthion	Ec	The lateral corner (angle) of the eye.	Pg. 07
2	Endocanthion	En	The medial angle of the eye. Medial corner of the eye where the eyelids meet, not in the caruncles (reddish eminence in the medial region of the eye).	Pg. 10
3	Iridion laterale	Il	The most lateral point of the iris rim.	Pg. 13
4	Iridion mediale	Im	The most medial point of the iris rim.	Pg. 14
5	Glabella	G	The most prominent region in the midsagittal plane between supraorbital arches.	Pg. 65
6	Nasion	N	Median point at the nasal root (apex of the frontonasal angle).	Pg. 66
7	Subnasal	Sn	Midpoint of the base of the columella, underneath the nasal spine.	Pg. 33
8	Alare	Al	The most lateral point of the nose wing. The most lateral point of the curvature of the nasal wing.	Pg. 35
9	Chelion	Ch	The corner of the mouth. The region of encounter of upper and lower lip vermilion border.	Pg. 42
10	Labiale superius	Ls	The midpoint (at the midsagittal plane) of the upper lip vermilion border.	Pg. 40
11	Stomion	Sto	The encounter of upper and lower lip at the midsagittal plane when lips are naturally closed.	Pg. 46
12	Labiale inferius	Li	The midpoint (at the midsagittal plane) of the lower lip vermilion border.	Pg. 47
13	Labiomentale	Lm	Point of greatest depression between the lower lip and the menton (at the mentolabial sulcus).	Pg. 48
14	Gnathion	Gn	The lowest point of menton edge, at the midsagittal plane.	Pg. 49
15	Gonion	Go	The most lateral point of the mandible angle. The widest point of the mandible.	Pg. 50
16	Zygion	Zy	The most lateral point (greater width) of the zygomatic bone (cheek).	Pg. 52

PHOTO-ANTHROPOMETRIC ANALYSIS: MANUAL LANDMARKING
07

1. ECTOCANTHION

Number	Landmark name	Laterality	Abbreviation
1	Ectocanthion	Bilateral	Ec_R / Ec_L

Photo-anthropometric definition

Most lateral landmark in the corner of the eye (distant from the midline), where the upper and lower ciliary implantation lines meet.

SAFF-2D landmarking procedures

Image approximation (Zoom):
Framing both eyes or the examined one specifically.

References:
Ciliary implantation lines.

Auxiliary lines:
Vertical and horizontal.

Procedure:
Move the vertical line from lateral to medial side of the face to the landmark where upper and lower ciliary lines meet, in the region of lateral angle of the eye. Then move the horizontal line until the point of convergence of those lines. Mark Ectocanthion in the intersection region between the two auxiliary lines. Follow the same procedure for landmarking the contralateral one.

FIG. 2—Example description for the Ectocanthion landmark taken from the facial photo-anthropometric (FPA) manual used in this work for the adapted positioning approach (AdMet) and provided in the Appendix S1. [Color figure can be viewed at wileyonlinelibrary.com]

intra-examiner scores were reliable (ICC > 0.75) for both EG and NG (i.e., the groups of expert and non-expert examiners, respectively) using both tested FPA approaches, that is, CIMet and AdMet.

Application of CIMet

Location dispersions for the 16 landmarks were calculated for both positioning approaches and were reported in Table 2

TABLE 2—Summary dispersion statistics (in px) for the 16 investigated landmarks according to the group of examiners (EG vs. NG) and the applied landmark-positioning approach (CIMet vs. AdMet)

Landmark	D	CIMet						AdMet					
		EG			NG			EG			NG		
		Mean	SD	Rank	Mean	SD	Rank	Mean	SD	Rank	Mean	SD	Rank
Al	D _X	0.812	0.731	14	1.057	0.871	16	0.548	0.440	14	0.681	0.602	14
	D _Y	2.411	1.859	7	3.502	2.722	8	1.440	1.161	7	1.588	1.530	7
	D _{MX_Y}	1.609	1.027	9	2.281	1.534	13	1.031	0.651	12	1.126	0.847	11
Ch	D _X	1.842	1.455	8	4.590	4.017	5	3.182	2.238	1	2.313	1.946	1
	D _Y	0.637	0.536	16	1.056	0.831	15	1.018	0.842	11	0.678	0.588	14
	D _{MX_Y}	1.239	0.812	11	2.832	2.113	10	2.102	1.255	3	1.521	1.031	5
Ec	D _X	3.858	2.433	2	4.166	2.676	6	2.088	1.784	3	2.041	1.587	3
	D _Y	1.321	0.938	9	1.678	1.217	13	1.340	1.017	8	1.887	1.611	4
	D _{MX_Y}	2.593	1.422	7	2.922	1.503	9	1.721	1.128	6	1.966	1.212	3
En	D _X	1.304	1.061	12	2.286	1.930	12	1.366	1.304	8	1.790	1.621	4
	D _Y	0.860	0.704	14	1.144	0.855	14	1.024	0.773	10	1.069	1.037	9
	D _{MX_Y}	1.077	0.688	14	1.723	1.222	15	1.202	0.826	11	1.426	1.207	7
G	D _X	2.839	2.492	3	2.718	1.856	10	1.267	1.077	9	0.762	0.806	12
	D _Y	14.423	8.488	2	75.721	18.755	1	1.921	1.521	4	2.033	1.887	3
	D _{MX_Y}	8.632	4.751	2	39.221	9.611	1	1.600	1.028	8	1.385	1.061	8
Gn	D _X	2.066	1.570	5	3.133	2.202	8	2.214	2.170	2	1.111	1.780	9
	D _Y	1.087	0.847	11	2.431	2.911	11	1.477	2.512	6	1.750	2.377	5
	D _{MX_Y}	1.579	0.930	10	2.758	1.777	11	1.853	1.611	5	1.453	1.437	6
Go	D _X	6.403	4.526	1	7.970	5.682	4	1.018	1.030	13	0.721	0.555	13
	D _Y	14.650	11.382	1	20.869	14.567	4	0.820	0.693	14	0.856	0.627	13
	D _{MX_Y}	10.517	7.804	1	14.417	9.877	4	0.917	0.646	14	0.775	0.479	13
Il	D _X	0.617	0.543	16	10.543	8.670	2	0.491	0.381	15	0.522	0.441	15
	D _Y	1.059	0.782	12	2.674	3.732	10	0.982	0.770	12	1.051	0.943	10
	D _{MX_Y}	0.838	0.480	16	6.606	4.579	5	0.727	0.433	15	0.771	0.564	14
Im	D _X	0.804	1.672	15	8.427	6.073	3	0.449	0.371	16	0.460	0.317	16
	D _Y	1.210	1.631	10	1.739	1.363	12	0.946	0.813	13	0.968	0.866	12
	D _{MX_Y}	1.032	1.586	15	5.093	3.013	6	0.702	0.458	16	0.703	0.455	15
Li	D _X	1.722	1.289	9	2.330	1.820	11	1.730	1.366	5	1.040	1.122	10
	D _Y	1.527	1.414	8	7.676	6.637	5	1.627	1.624	5	1.342	1.226	8
	D _{MX_Y}	1.627	0.991	8	5.011	3.403	7	1.681	1.077	7	1.185	0.789	9
Lm	D _X	1.952	1.728	6	3.069	2.179	9	1.766	1.232	4	1.155	1.304	7
	D _Y	6.465	5.030	5	5.548	4.252	6	3.864	4.845	2	5.568	9.725	2
	D _{MX_Y}	4.208	2.692	5	4.313	2.423	8	2.816	2.563	2	3.371	4.828	2
Ls	D _X	1.579	1.111	11	1.430	1.126	14	1.268	0.923	10	1.221	0.951	6
	D _Y	4.111	3.614	6	3.524	3.140	7	2.688	2.414	3	1.047	0.830	11
	D _{MX_Y}	2.838	2.017	6	2.478	1.588	12	1.982	1.376	4	1.131	0.673	10
N	D _X	1.876	1.322	7	3.141	2.363	7	1.220	1.070	11	0.786	0.792	11
	D _Y	7.542	6.868	4	58.842	50.286	2	0.709	0.583	16	0.615	0.660	15
	D _{MX_Y}	4.706	3.603	4	30.989	25.664	2	0.955	0.555	13	0.729	0.512	16
Sn	D _X	1.209	0.930	13	1.210	0.851	15	1.611	1.210	7	1.380	1.285	5
	D _Y	0.976	0.848	13	2.943	4.945	9	1.245	1.121	9	1.742	1.559	6
	D _{MX_Y}	1.101	0.662	13	2.067	2.504	14	1.416	0.857	9	1.555	1.033	4
Sto	D _X	1.640	1.211	10	1.869	1.548	13	1.727	1.280	6	1.133	1.137	8
	D _Y	0.711	0.683	15	0.981	0.931	16	0.710	0.466	15	0.611	0.488	16
	D _{MX_Y}	1.185	0.727	12	1.428	0.890	16	1.218	0.702	10	0.866	0.676	12
Zy	D _X	2.588	2.072	4	23.311	13.214	1	1.061	0.832	12	2.094	2.065	2
	D _Y	11.656	8.201	3	21.532	13.007	3	6.833	6.023	1	14.122	12.127	1
	D _{MX_Y}	7.117	4.327	3	22.423	10.628	3	3.939	3.051	1	8.104	6.666	1
Global	D _X	2.069	-	-	5.078	-	-	1.438	-	-	1.201	-	-
	D _Y	4.415	-	-	13.241	-	-	1.790	-	-	2.308	-	-
	D _{MX_Y}	3.244	-	-	9.160	-	-	1.616	-	-	1.754	-	-

D, dispersion statistics; SD, standard deviation.

(values in px, i.e., D_X, D_Y, and D_{MX_Y}), and Table 3 (values converted in mm, i.e., ERD_X, ERD_Y, and ERD_{MX_Y}). A graphical comparison of D_{MX_Y} and the analysis of effects are furthermore displayed in Figs 3 and 4, respectively.

Using CIMet, the two groups of examiners performed the landmark-positionings very differently, with EG showing significantly better results than NG. In fact, the mean D_{MX_Y} values were 3.244 px (0.746 mm) and 9.160 px (2.106 mm) for NG and EG, respectively, which corresponds to a difference >2.8 times. The highest D_{MX_Y} for NG (i.e., 39.221 px or 9.016 mm

for G) was almost 4 times larger than the highest D_{MX_Y} for EG (i.e., 10.517 px or 2.418 mm for Go). Furthermore, 12 of the 16 landmarks (Al, Ch, En, G, Gn, Go, Il, Im, Li, N, Sn, and Zy) were significantly more dispersed for NG than for EG. Consequently, positioning performances with CIMet were proved to be strongly dependent on the previous anatomical knowledge and/or experience of the examiners, with more experienced examiners providing significantly more reproducible results.

More generally, Go, G, Zy, and N showed the largest dispersions in both groups of examiners and were thus the most

TABLE 3—Summary dispersion statistics (after conversion to mm) for the 16 investigated landmarks according to the group of examiners (EG vs. NG) and the applied landmark-positioning approach (CIMet vs. AdMet). A comparison of the values with reference clinical thresholds previously reported in the literature is also given in the columns headed “T.”

Landmark	ERD	CIMet						AdMet					
		EG			NG			EG			NG		
		Mean	SD	T.	Mean	SD	T.	Mean	SD	T.	Mean	SD	T.
Al	ERD _X	0.187	0.168		0.243	0.200		0.126	0.101		0.157	0.138	
	ERD _Y	0.554	0.427		0.805	0.626	x	0.331	0.267		0.365	0.352	
	ERD _{MXY}	0.370	0.236		0.524	0.353		0.237	0.150		0.259	0.195	
Ch	ERD _X	0.423	0.334		1.055	0.923	xx	0.731	0.514	x	0.532	0.447	
	ERD _Y	0.146	0.123		0.243	0.191		0.234	0.194		0.156	0.135	
	ERD _{MXY}	0.285	0.187		0.651	0.486	x	0.483	0.289		0.350	0.237	
Ec	ERD _X	0.887	0.559	x	0.958	0.615	x	0.480	0.410		0.469	0.365	
	ERD _Y	0.304	0.216		0.386	0.280		0.308	0.234		0.434	0.370	
	ERD _{MXY}	0.596	0.327	x	0.672	0.346	x	0.396	0.259		0.452	0.279	
En	ERD _X	0.300	0.244		0.526	0.444		0.314	0.300		0.411	0.373	
	ERD _Y	0.198	0.162		0.263	0.197		0.235	0.178		0.246	0.238	
	ERD _{MXY}	0.248	0.158		0.396	0.281		0.276	0.190		0.328	0.277	
G	ERD _X	0.653	0.573	x	0.625	0.427	x	0.291	0.248		0.175	0.185	
	ERD _Y	3.316	1.951	xx	17.407	4.311	xx	0.442	0.350		0.467	0.434	
	ERD _{MXY}	1.984	1.092	xx	9.016	2.209	xx	0.368	0.236		0.318	0.244	
Gn	ERD _X	0.475	0.361		0.720	0.506	x	0.509	0.499		0.255	0.409	
	ERD _Y	0.250	0.195		0.559	0.669		0.340	0.577		0.402	0.546	
	ERD _{MXY}	0.363	0.214		0.634	0.409	x	0.426	0.370		0.334	0.330	
Go	ERD _X	1.472	1.040	xx	1.832	1.306	xx	0.234	0.237		0.166	0.128	
	ERD _Y	3.368	2.617	xx	4.797	3.349	xx	0.189	0.159		0.197	0.144	
	ERD _{MXY}	2.418	1.794	xx	3.314	2.271	xx	0.211	0.149		0.178	0.110	
Il	ERD _X	0.142	0.125		2.424	1.993	xx	0.113	0.088		0.120	0.101	
	ERD _Y	0.243	0.180		0.615	0.858	x	0.226	0.177		0.242	0.217	
	ERD _{MXY}	0.193	0.110		1.519	1.053	xx	0.167	0.100		0.177	0.130	
Im	ERD _X	0.185	0.384		1.937	1.396	xx	0.103	0.085		0.106	0.073	
	ERD _Y	0.278	0.375		0.400	0.313		0.217	0.187		0.223	0.199	
	ERD _{MXY}	0.237	0.365		1.171	0.693	xx	0.161	0.105		0.162	0.105	
Li	ERD _X	0.396	0.296		0.536	0.418		0.398	0.314		0.239	0.258	
	ERD _Y	0.351	0.325		1.765	1.526	xx	0.374	0.373		0.309	0.282	
	ERD _{MXY}	0.374	0.228		1.152	0.782	xx	0.386	0.248		0.272	0.181	
Lm	ERD _X	0.449	0.397		0.706	0.501	x	0.406	0.283		0.266	0.300	
	ERD _Y	1.486	1.156	xx	1.275	0.977	xx	0.888	1.114	x	1.280	2.236	xx
	ERD _{MXY}	0.967	0.619	x	0.991	0.557	x	0.647	0.589	x	0.775	1.110	x
Ls	ERD _X	0.363	0.255		0.329	0.259		0.291	0.212		0.281	0.219	
	ERD _Y	0.945	0.831	x	0.810	0.722	x	0.618	0.555	x	0.241	0.191	
	ERD _{MXY}	0.652	0.464	x	0.570	0.365		0.456	0.316		0.260	0.155	
N	ERD _X	0.431	0.304		0.722	0.543	x	0.280	0.246		0.181	0.182	
	ERD _Y	1.734	1.579	xx	13.527	11.560	xx	0.163	0.134		0.141	0.152	
	ERD _{MXY}	1.082	0.828	xx	7.124	5.900	xx	0.220	0.128		0.168	0.118	
Sn	ERD _X	0.278	0.214		0.278	0.196		0.370	0.278		0.317	0.295	
	ERD _Y	0.224	0.195		0.677	1.137	x	0.286	0.258		0.400	0.358	
	ERD _{MXY}	0.253	0.152		0.475	0.576		0.326	0.197		0.357	0.237	
Sto	ERD _X	0.377	0.278		0.430	0.356		0.397	0.294		0.260	0.261	
	ERD _Y	0.163	0.157		0.226	0.214		0.163	0.107		0.140	0.112	
	ERD _{MXY}	0.272	0.167		0.328	0.205		0.280	0.161		0.199	0.155	
Zy	ERD _X	0.595	0.476	x	5.359	3.038	xx	0.244	0.191		0.481	0.475	
	ERD _Y	2.680	1.885	xx	4.950	2.990	xx	1.571	1.385	xx	3.246	2.788	xx
	ERD _{MXY}	1.636	0.995	xx	5.155	2.443	xx	0.906	0.701	x	1.863	1.532	xx
Global	ERD _X	0.476	-		1.167	-	xx	0.331	-		0.276	-	
	ERD _Y	1.015	-	xx	3.044	-	xx	0.412	-		0.531	-	
	ERD _{MXY}	0.746	-	x	2.106	-	xx	0.372	-		0.403	-	

For thresholds: “xx” = above acceptable limits (>1 mm), “x” = within the range of acceptability (0.575 and 1 mm), values without crosses were within an ideal average dispersion (<0.575 mm). ERD, estimated real mean dispersion; SD, standard deviation; T., reference threshold.

difficult landmarks to positioning. On the contrary, En, Sn, and Sto were generally within the five least dispersed landmarks overall.

Adoption of AdMet

Adoption of AdMet resulted in a significant decrease in the dispersion of landmark placement for both groups of examiners. This was particularly true for NG. Indeed, its mean D_{MXY}

passed from 9.160 px (2.106 mm) to 1.754 px (0.403 mm), compared to a decrease from 3.244 px (0.746 mm) to 1.616 px (0.372 mm) for EG. A statistically significant decrease was furthermore observed in the positioning dispersion of 13 of the 16 landmarks (Al, Ch, Ec, G, Gn, Go, Il, Im, Li, Ls, N, Sto, and Zy) for NG, and in that of 10 of the 16 landmarks (Al, Ec, G, Go, Il, Im, Lm, Ls, N, and Zy) for EG. These results together proved that the use of AdMet actually significantly improved reproducibility in landmark positioning, independent from

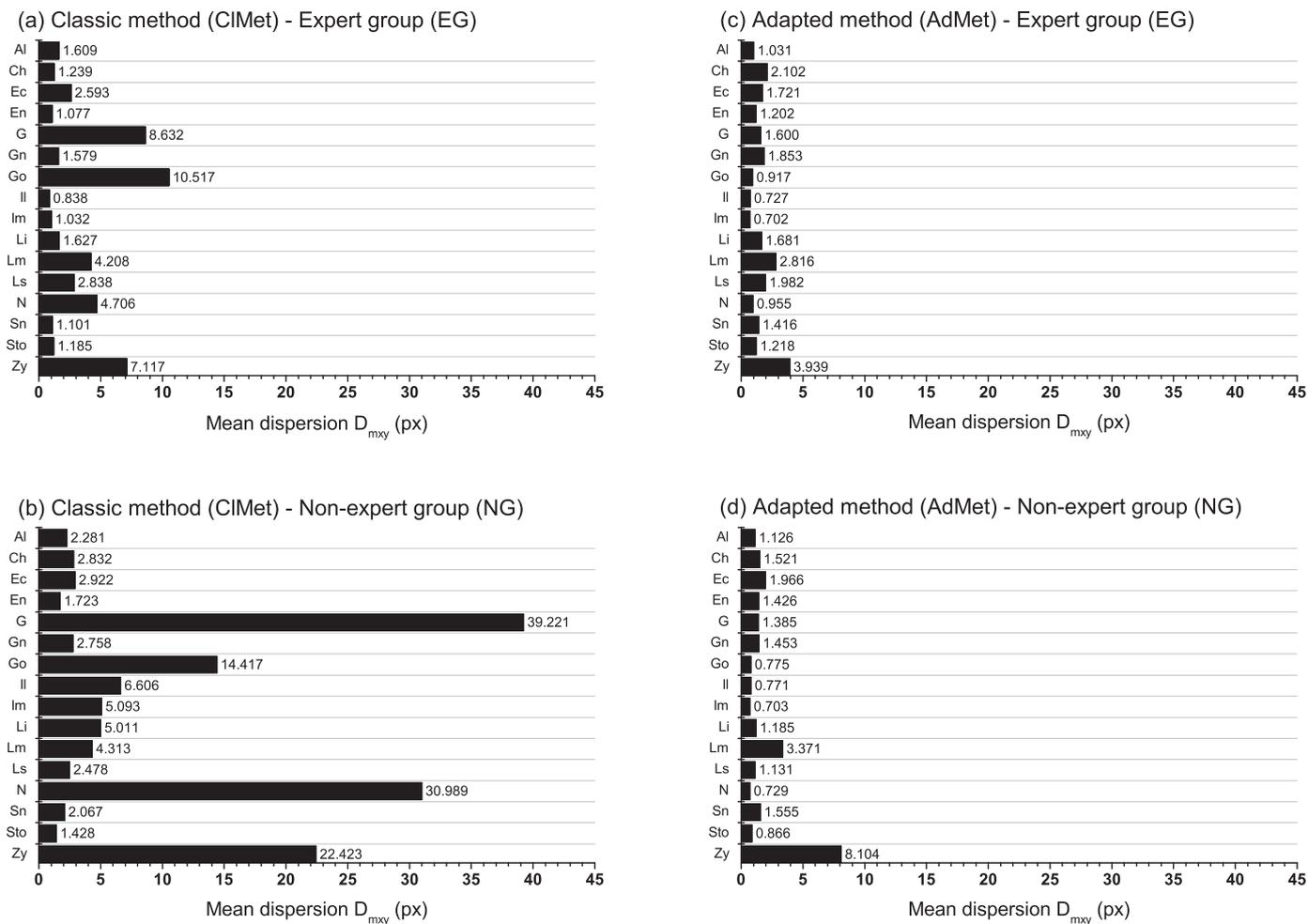


FIG. 3—Comparison of the mean intra-landmark dispersion values (D_{MXY}) observed in the positioning of the 16 landmarks using the different experimental settings.

previous anatomical knowledge and/or experience of the examiner. A simultaneous increase in the positioning dispersion of two of the 16 landmarks (Sn and Ch) was, nevertheless, detected for EG. Even if statistically significant, however, this was still really small on a physical scale and thus considered negligible from a practical point of view (Fig. 4).

Comparison of the results obtained by the two groups of examiners between themselves showed that, on average, they performed very similarly when the novel landmark-positioning approach was used. In fact, the respective mean D_{MXY} values were largely consistent (1.616 px or 0.372 mm for EG, and 1.754 px or 0.403 mm for NG). Perhaps surprising, however, was that dispersion results for the single landmarks showed that, from a statistical point of view, a higher number of landmarks were more reproducibly positioned by NG compared to EG. Indeed, seven over 16 landmarks (Ch, Gn, Go, Li, Ls, N, and Sto) showed significantly lower D_{MXY} values for NG than for EG when AdMet was used, while only one of 16 (Zy) showed a significantly larger D_{MXY} . Again, the differences in dispersion for all these landmarks (but Zy) were very small on a physical scale and, thus, considered inconsequential from a practical point of view (Fig. 4). Hence, it could be concluded that AdMet allowed for a higher degree of robustness in landmark-positioning between examiners with different anatomical knowledge and/or experience. The only exception was the placement of Zy, for which previous knowledge and/or experience seemed particularly important.

More generally, the positioning of Zy resulted in relatively high D_{MXY} values for both groups of examiners, especially when compared to the other landmarks. This was particularly true for NG, as the D_{MXY} for this landmark was 8.104 px (1.863 mm) against 3.939 px (0.906 mm) for EG. The dispersions of the other three landmarks that showed particularly high D_{MXY} using CIMet (i.e., G, N, and Go) were significantly decreased through the use of AdMet.

Clinical Validation

In order to validate the approach against clinically accepted standards, estimated real mean dispersion (ERD_{MXY}) values were compared to reference thresholds previously reported in the cephalometric literature (Table 4). A complete comparison for all ERD values (i.e., ERD_x , ERD_y , and ERD_{MXY}) is further available in Table 3.

CIMet led to ERD_{MXY} within ideal or acceptable limits (i.e., ≤ 1 mm) for several landmarks when used by both groups of examiners (12 landmarks for EG and nine for NG). A significant number of landmarks, however, showed ERD_{MXY} above acceptable limits (i.e., >1 mm); these were, namely, four landmarks for EG (i.e., G, Go, N, and Zy) and seven for NG (i.e., G, Go, Il, Im, Li, N, and Zy). For NG, in particular, G, Go, N, and Zy showed ERD_{MXY} larger than 3 mm, which were considered especially high. When AdMet was used, none of the landmarks



FIG. 4—Graphical comparison of the mean intra-landmark dispersion values (D_{MXY}) observed in the positioning of the 16 landmarks when different experimental settings were adopted (landmark-positioning approaches on left; examiners on right). The columns “Var.” (variability) visually represent the overlap of the dispersions considering a 50-pixel scale. The columns “Sig.” (significance), on the contrary, represent the statistical significance of the dispersion differences ($\alpha = 0.05$) using a color scale. [Color figure can be viewed at wileyonlinelibrary.com]

showed ERD_{MXY} above acceptable limits for either group of examiners, except one for NG (i.e., Zy). More specifically, 14 landmarks were generally within the ideal range (i.e., <0.575 mm). This showed the higher validity of AdMet when compared with previously reported clinical standards.

Discussion

Physical anthropology is a well-established tool for the extraction, interpretation, and classification of the human body within industrial, medical, orthodontic, and forensic applications (14,25,30,31). In recent decades, the increasingly widespread use of digital imaging devices has highlighted the necessity of bringing its precepts to

indirect, 2D-image contexts. Starting from the assumption that all FPA-based analyses (e.g., establishment of measures, angles, ratios, and indexes) rely on the previous determination of landmarks, evaluating the particular variation regarding their positioning is a necessary step for its safe and reliable application (1,2,12,13).

Although landmark-positioning variability has been a commonly addressed issue in the scientific community, its assessment and improvement for uses on photographs have been scarce. In particular, no studies have ever proposed conceptual adaptations to the definition of landmarks for image-based applications, while those that have addressed FPA-positioning variability used non-specific landmark-positioning approaches (i.e., cephalometric definitions). As a consequence, doubts can be raised concerning the

TABLE 4—Comparison of the estimated real mean dispersions (ERD_{MXY}) with reference clinical thresholds previously reported in the literature

Landmark	CIMet				AdMet			
	EG		NG		EG		NG	
	ERD_{MXY}	Thres.	ERD_{MXY}	Thres.	ERD_{MXY}	Thres.	ERD_{MXY}	Thres.
Al	0.370		0.524		0.237		0.259	
Ch	0.285		0.651	x	0.483		0.350	
Ec	0.595	x	0.672	x	0.396		0.452	
En	0.248		0.396		0.276		0.328	
G	1.984	xx	9.016	xx	0.368		0.318	
Gn	0.363		0.634	x	0.426		0.334	
Go	2.418	xx	3.314	xx	0.211		0.178	
Il	0.193		1.519	xx	0.167		0.177	
Im	0.237		1.171	xx	0.161		0.162	
Li	0.374		1.152	xx	0.386		0.272	
Lm	0.967	x	0.991	x	0.647	x	0.775	x
Ls	0.652	x	0.570		0.456		0.260	
N	1.082	xx	7.124	xx	0.220		0.168	
Sn	0.253		0.475		0.326		0.357	
Sto	0.272		0.328		0.280		0.199	
Zy	1.636	xx	5.155	xx	0.906	x	1.863	xx
Global	0.746	x	2.106	xx	0.372		0.403	

“xx” : above acceptable limits (>1 mm); “x” : within the range of acceptability (0.575 and 1 mm); values without crosses have an ideal average dispersion (<0.575 mm).

proper and reliable attribution of the investigated landmarks (6,17,19). Recently, an alternative nomenclature (i.e., capulometric landmarks) has been tentatively proposed for the analysis of 2D images (22). Again, nonetheless, no visual references were implemented, resulting in a set of definitions very similar to the classic cephalometric ones. The lack of a standardized set of landmarks and protocols specific to FPA analysis should be viewed with concern because, depending on the scientific field of interest, errors may lead to misunderstandings in diagnosis/treatment or even to improper characterization and/or classification of a specific population or individual (3,7).

Classifying human features into class or individual characteristic is a constant practice in forensic science. A proper population survey of a specific facial feature, whether morphological or photo-anthropometrical, is necessary to determine its importance in the human individualization process and to statistically support the quantification and decision of an identification match (20,39,40). As a result of its inherent potential to make image-based facial analysis more objective, systematic, and reproducible, FPA has promising capabilities for the analytical survey of facial structures along with the high possibility of automatization. This is a step forward for the evaluation of large databases, as well for understanding human facial variation. In this sense, generating landmark-specific variability information according to the adopted methodology is of utmost importance, by determining the extent to which each one can provide reliable facial relationships to support forthcoming statistical associations.

In the present study, as expected, the use of classical cephalometric descriptions led to low reproducibilities between the examiners in positioning the 16 investigated landmarks on facial images. Indeed, ERD_{MXY} values for most of them were above an ideal limit threshold, and this was true not only for non-expert examiners, but also for expert ones. More specifically, only nine of the 16 landmarks showed ERD_{MXY} values within an ideal error range when positioned by expert examiners, and four of 16 had ERD_{MXY} values above an acceptable threshold.

Observed dispersions, furthermore, showed an overall low consistency between the two groups of examiners, with

non-experts particularly struggling with placing landmarks on facial images in a reproducible way, as demonstrated by their significantly bigger inter-variability. This suggests a low robustness of the classic landmark-positioning method with respect to the experience level of the examiners and, in particular, that previous anatomical knowledge and/or experience in the procedure are necessary in order to properly understand traditional cephalometric descriptions and locate the corresponding structures on facial images.

The positioning of Go, G, Zy, and N on frontal facial images proved to be particularly challenging following the traditional cephalometric descriptions, as proved by their very high dispersions amongst all the examiners (especially non-experts). This is a serious problem that may affect the usefulness of the traditional landmark method in many FPA applications. Indeed, these four specific landmarks are involved in the establishment of some of the most characteristic facial measurements (14,29), such as the facial height (N - Gn), facial width (Zy - Zy), mandibular width (Go - Go), facial length index (N - Gn/Zy - Zy), mandibulo-facial index (Go - Go/Zy - Zy), and naso-chelion angle (Ch - N - Ch). The same observation has, nonetheless, already been reported in a number of previous studies (6,17,22,35,41,42) and may be explained by the fact that the traditional cephalometric descriptions for these four landmarks largely rely on physical and/or bone structures, which are particularly difficult to detect on frontal images. As a proof, the opposite trend could actually be seen for landmarks such as Ch and Sto, for which traditional cephalometric definitions rely more strongly on facial structures visible on images (6,22).

The adoption of adapted and FPA-specific landmark definitions positively enhanced the performance of positioning the 16 investigated landmarks on facial images and, thus, of the general FPA procedure. Undeniably, placement reproducibility between examiners was significantly improved. All the landmarks showed ERD_{MXY} within acceptable limit thresholds when placed by expert examiners, contrary to that observed when classic cephalometric definitions were used. Even more notably, 14 of

16 landmarks showed ERD_{MXY} values within ideal limit thresholds. In contrast, landmark dispersions showed a better consistency between experts and non-experts. This finding supports the higher robustness of the adapted landmark approach with respect to the experience level of the examiners. Furthermore, it is also consistent with the conclusion that the most relevant factor in the correct positioning of landmarks on facial images is not necessarily the examiner's previous knowledge in facial anatomy or their experience in the procedure, but rather the accuracy of the landmark descriptions themselves. In this regard, an FPA-optimized approach is more helpful than a cephalometry-based one, as the latter is essentially based on descriptions of underlying anatomical structures.

The use of adapted landmark definitions also solved the high positioning variability of G, N, and Go that is observed when using the classic cephalometric approach; an improvement that, by itself, is prone to significantly enhance the general reliability of FPA in most applications. Placement of Zy, however, still resulted in high ERD_{MXY} for both groups of examiners, which confirms its particular complexity in being positioned on facial images. Nonetheless, after a more detailed inspection, it can be observed that its dispersion on the vertical axis (ERD_Y) more significantly contributes to ERD_{MXY} than its dispersion on the horizontal axis (ERD_X), and that the latter is almost negligible and within an ideal threshold after using an adapted landmark-positioning approach. In this regard, it is important to highlight that errors in the vertical and horizontal directions may be of substantial importance depending on the specific application and/or landmark. Zy, in particular, is most frequently used in horizontal measurements (e.g., facial width) and related indexes (e.g., facial length index) (14,29), and thus the use of an adapted approach may actually allow a more efficient use of this landmark. In any case, further improvements to the landmark descriptions may be implemented in order to also take into account the variability on the vertical axis and bring ERD_Y to within an acceptable dispersion range.

Conclusion

In this work, the use of an adapted approach for landmark facial images based on descriptions and locating procedures optimized for FPA analysis has been assessed and compared against a traditional approach based on classic cephalometric descriptions. Results showed that the use of conventional cephalometric descriptions led to a low reproducibility between examiners in positioning landmarks and, more importantly, to a low consistency in the positioning dispersions between experts and non-experts. This suggested that previous anatomical knowledge and/or experience is necessary in order to correctly apply traditional cephalometric descriptions. The use of adapted landmark definitions, on the contrary, significantly decreased the landmark dispersion between examiners, whilst also reducing the differences arising from experience level. This second observation, in particular, supported the conclusion that the most relevant factor in the correct positioning of landmarks on facial images is not necessarily the examiner's knowledge about facial anatomy, but instead the accuracy of landmark descriptions and the application of an approach based on clear visual references.

Thus, the use of an adapted landmark-positioning approach proved to be highly advantageous in FPA analysis and future work in this field should consider adopting similar methodologies. In particular, the adapted approach specifically used in this research performed well and may be implemented in future FPA applications.

Acknowledgments

The authors express their gratitude to the Technical and Scientific Section of the Brazilian Federal Police, especially the Unit of Forensic Electronics and Multimedia of the National Institute of Criminalistics, as well as to Edmar Antônio da Silva for the development of SAFF-2D[®] (Forensic Facial Analysis System—2D, Department of Federal Police, Brazil). The authors would also like to thank Rachel Irlam (King's College London, U.K.) and David Jedd (Birmingham, UK) for proofreading the document, as well as the Coordination of Improvement of Higher Level Personnel, Ministry of Education (CAPES Foundation-Brazil) and the Forensic Craniofacial Identification Scientific Group—Brazil (FACIS Group).

References

- Davis J, Valentine T, Davis R. Computer assisted photo-anthropometric analyses of full-face and profile facial images. *Forensic Sci Int* 2010;200(1–3):165–76.
- İşcan M. Introduction to techniques for photographic comparison: potentials and problems. In: İşcan M, Helmer R, editors. *Forensic analysis of the skull: craniofacial analysis, reconstruction, and identification*. New York, NY: Wiley-Liss, 1993;57–70.
- Stavrianos C, Papadopoulos C, Pantelidou O, Emmanouil J, Pentalotis N, Tatsis D. The use of photoanthropometry in facial mapping. *Res J Med Sci* 2012;6(4):166–9.
- Moreton R, Morley J. Investigation into the use of photoanthropometry in facial image comparison. *Forensic Sci Int* 2011;212(1–3):231–7.
- Borges DL, Vidal FB, Flores MRP, Melani RFH, Guimarães MA, Machado CEP. Photoanthropometric face iridial proportions for age estimation: an investigation using features selected via a joint mutual information criterion. *Forensic Sci Int* 2018;284:9–14.
- Cattaneo C, Obertová Z, Ratnayake M, Marasciulo L, Tutkuvieni J, Poppa P, et al. Can facial proportions taken from images be of use for ageing in cases of suspected child pornography? A pilot study. *Int J Legal Med* 2012;126(1):139–44.
- Machado CEP, Flores MRP, Lima LNC, Tinoco RLR, Franco A, Bezerra ACB, et al. A new approach for the analysis of facial growth and age estimation: iris ratio. *PLoS ONE* 2017;12(7):e0180330.
- Adamu L, Ojo S, Danborn B, Adebisi S, Taura M. Sex determination using facial linear dimensions and angles among Hausa population of Kano State, Nigeria. *Egypt J Forensic Sci* 2016;6(4):459–67.
- Packiriswamy V, Bashour M, Nayak S. Anthropometric analysis of the South Indian Woman's nose. *Facial Plast Surg* 2016;32(3):304–8.
- Ramanathan N, Chellappa R. Modeling age progression in young faces. *IEEE Proceedings of the 2006 Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*; 2006 June 17–22; New York, NY. Piscataway, NJ: Institute of Electrical and Electronics Engineers, 2006;387–94.
- Porter G, Doran G. An anatomical and photographic technique for forensic facial identification. *Forensic Sci Int* 2000;114(2):97–105.
- Halberstein RA. The application of anthropometric indices in forensic photography: three case studies. *J Forensic Sci* 2001;46(6):1438–41.
- Trpkova B, Major P, Prasad N, Nebbe B. Cephalometric landmarks identification and reproducibility: a meta analysis. *Am J Orthod Dentofacial Orthop* 1997;112(2):165–70.
- Farkas L. *Anthropometry of the head and face*, 2nd edn. New York, NY: Raven Press, 1994.
- Shishkin KM, Arsenina OI, Shishkin MK, Popova NV. Cephalometry efficacy in orthodontic treatment planning: correlations of cephalometric values and their changes in the course of treatment. *Stomatologia* 2017;96(4):36–7.
- Esenlik E, Plana NM, Grayson BH, Flores RL. Cephalometric predictors of clinical severity in Treacher Collins syndrome. *Plast Reconstr Surg* 2017;140(6):1240–9.
- Campomanes-Álvarez B, Ibáñez O, Navarro F, Alemán I, Córdón O, Damas S. Dispersion assessment in the location of facial landmarks on photographs. *Int J Legal Med* 2015;129(1):227–36.
- Cattaneo C, Ritz-Timme S, Gabriel P, Gibelli D, Giudici E, Poppa P, et al. The difficult issue of age assessment on pedo-pornographic material. *Forensic Sci Int* 2009;183(1–3):21–4.
- Lucas T, Kumaratilake J, Henneber M. Metric identification of the same people from images: how reliable is it? *J Anthropol* 2016;2016(1):1–10.

20. Cummaudo M, Guerzoni M, Marasciuolo L, Gibelli D, Cigada A, Oberrovà Z, et al. Pitfalls at the root of facial assessment on photographs: a quantitative study of accuracy in positioning facial landmarks. *Int J Legal Med* 2013;127(3):699–706.
21. FISWG. Guidelines for facial comparison methods; www.fiswg.org (accessed July 22, 2018).
22. Caple J, Stephan CN. A standardized nomenclature for craniofacial and facial anthropometry. *Int J Legal Med* 2016;130(3):863–79.
23. Lee WJ, Kim DM, Lee UY, Cho JH, Kim MS, Hong JH, et al. A preliminary study of the reliability of anatomical facial landmarks used in facial comparison. *J Forensic Sci* 2018;1–9. <https://doi.org/10.1111/1556-4029.13873>. Epub 2018 Aug 14.
24. Wilkinson C, Evans R. Are facial image analysis experts any better than the general public at identifying individuals from CCTV images? *Sci Justice* 2009;49(3):191–6.
25. Farkas L. Accuracy of anthropometric measurements: past, present, and future. *Cleft Palate Craniofac J* 1996;33(1):10–8; discussion 19–22.
26. Driessen J, Vuyk H, Borgstein J. New insights into facial anthropometry in digital photographs using iris dependent calibration. *Int J Pediatr Otorhinolaryngol* 2011;75(4):579–84.
27. Miot H, Pivotto D, Jorge E, Mazeto G. Evaluation of oculometric parameters by facial digital photography: use of iris diameter as a reference. *Arq Bras Oftalmol* 2008;71(5):679–83.
28. Flores MRP, Machado CEP, Silva RHA. Proposta de análise facial fotoantropométrica em norma frontal: metodologia descritiva dos pontos anatômicos de referência [Photo-anthropometric facial analysis in frontal view images: descriptive methodology proposal for anatomical landmarks]. *Novas Edições Acadêmicas*, 2017; http://facisgroup.org/facial_landmarks (accessed August 11, 2018).
29. George R. Facial geometry: graphic facial analysis for forensic artists. Springfield, MO: Charles C Thomas Publisher, 2007.
30. Kolar J, Salter E. Craniofacial anthropometry: practical measurement of the head and face for clinical, surgical, and research use. Springfield, MO: Charles C Thomas Publisher, 1997.
31. Zimble M, Ham J. Aesthetic facial analysis. In: Cummings CW, Flint PW, Harker LA, et al., editors. *Cummings otolaryngology: head and neck surgery*, 4th edn. Philadelphia, PA: Elsevier Mosby, 2005;517–8.
32. Hickson-Curran S, Young G, Brennan N, Hunt C. Chinese and Caucasian ocular topography and soft contact lens fit. *Clin Exp Optom* 2016;99(2):149–56.
33. Ronneburger A, Basarab J, Howland H. Growth of the cornea from infancy to adolescence. *Ophthalmic Physiol Opt* 2006;26(1):80–7.
34. Martin D, Holden B. A new method for measuring the diameter of the in vivo human cornea. *Am J Optom Physiol Opt* 1982;59(5):436–41.
35. Aksu M, Kaya D, Kocadereli I. Reliability of reference distances used in photogrammetry. *Angle Orthod* 2010;80(4):482–9.
36. Forsyth D, Davis D. Assessment of an automated cephalometric analysis system. *Eur J Orthod* 1996;18(5):471–8.
37. Richardson A. A comparison of traditional and computerized methods of cephalometric analysis. *Eur J Orthod* 1981;3(1):15–20.
38. Strauss R, Weis B, Lindauer S, Rebellato J, Isaacson R. Variability of facial photographs for use in treatment planning for orthodontics and orthognathic surgery. *Int J Adult Orthodon Orthognath Surg* 1997;12(3):197–203.
39. Mallett XD, Dryden I, Bruegge RV, Evison M. An exploration of sample representativeness in anthropometric facial comparison. *J Forensic Sci* 2010;55(4):1025–31.
40. 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.
41. Bishara S, Cummins D, Jorgensen G, Jakobsen J. A computer assisted photogrammetric analysis of soft tissue changes after orthodontic treatment. Part I: methodology and reliability. *Am J Orthod Dentofacial Orthop* 1995;107(7):633–9.
42. Farkas L, Bryson W, Klotz J. Is photogrammetry of the face reliable? *Plast Reconstr Surg* 1980;66(3):346–55.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Manual of facial photo-anthropometry: visual references for landmark positioning in frontal view images.