p-ISSN: 2184-2205 e-ISSN: 2184-3414 ISBN: 978-989-35106-0-5 © 2023
https://doi.org/10.36315/2023inpact036

AN INTERDISCIPLINARY APPROACH TO THE CONSTRUCTION OF FACIAL COMPOSITES IN THE MEXICAN POPULATION: AI, ANTHROPOLOGY AND PSYCHOLOGY

Katya Rodríguez[1], Luis Fernando Cuevas[2], Sergio Padilla[1], Arodi Farrera[3], Germán Palafox[2], & Gustavo Ortiz[2]
[1]Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, UNAM (México)
[2]Facultad de Pscicología, UNAM (México)
[3]Instituto de Investigaciones Antropolóogicas, UNAM (México)

Abstract

We present an interdisciplinary approach to the construction of facial composites that could overcome some of the cognitive limitations that are frequently found in eyewitness/forensic recognition. We take advantage of the extensive anthropological research carried out on the facial characteristics of the Mexican population that produce the CARAMEX database, and combine it with the most recent methodologies of artificial intelligence and image processing to construct a large database of synthetic faces. Additionally, we present preliminary results on a corpus of basic linguistic/verbal facial descriptions obtained with subsamples of CARAMEX faces, to be used as the initial seed for the evolution and gradual convergence on a target facial composite.

Keywords: *Facial composite, holistic criterion, AI algorithms, anthropometry, Mexican population.*

1. Justification

Face recognition in forensic settings has ample space for improvement. The most common technique used by police departments around the world is the construction of facial composites, in which crime victims or eyewitnesses provide a verbal description of a suspect's facial features to a forensic art expert to deliver a graphic representation of the suspect's face. However, the ability to identify a person from a facial composite, even when the face is familiar, ranges from 5 to 20% (Bruce, et al. 2002; Frowd, et al. 2007; Klum, et al. 2013; Zahradnikova, et al. 2018); in part, this is probably due to the fact that the traditional way of constructing facial composites follows a feature-by-feature logic that by itself does not capture the complexities of human facial recognition. Thus, although it is possible to direct attention to individual features of a face, such as the eyes or the mouth, other facial features modify the perception of the focused feature in a holistic way (Meltzer & Bartlett 2019; Tanaka & Farah 1993; Tanaka & Sengco 1997). For example, it has been shown that it is much easier to identify a feature in a familiar face (e.g., a nose) if the feature is embedded in the context of a full face, than when the same feature is shown in isolation (Tanaka & Farah 1993; Tanaka & Simonyi 2016).

Also, the verbal descriptions that form the basic input for the construction of facial composites might interfere with the recognition of a briefly observed unfamiliar face by altering or modifying the mental representations of the face, a phenomenon that is known as verbal overshadowing (Baker & Reysen, 2020; Brown, Portch, Nelson & Frowd 2020; Meissner & Brigham, 2001).

In the last two decades, efforts to improve the construction of facial composites have led to the development of different semi-automatized systems that try to overcome the problems mentioned above (e.g., relying on a more holistic or "natural" way of perceiving faces; less dependence on verbal descriptions) using artificial intelligence and/or advanced image processing techniques (Frowd, 2021; Zahradnikova, et al. 2018). Unfortunately, forensic face recognition rates are still unsatisfactorily low.

In our own effort to come up with more efficient ways to create composite faces and improve forensic face recognition, we have used Artificial Intelligence (AI) image processing techniques and convolutional neural networks (CNN) on a representative sample of the facial variability of the Mexican population (see description of the CARAMEX database below), to create a more dense face space of real and synthetic faces that could serve as a search map for the recognition process.

Additionally, we collected linguistic/verbal descriptions of facial features from a sample of CARAMEX faces to explore ways to refine the search in the face space using natural language algorithms to standardize the descriptions. In parallel, the linguistic/verbal descriptions will be used to build a corpus of facial descriptors that could be used to facilitate a victim's or eyewitness' description of the face of a suspect with less overshadowing or induced memory distortions.

Thus, we present here both the results of the application of an AI image processing and CNN on CARAMEX faces to build a large database of synthetic faces, and some preliminary results of the linguistic/verbal facial descriptors that we plan to use to guide a search in the face space. For these two tasks, working with the CARAMEX database provided both an statistically and anthropologically relevant referent for facial recognition, given the sampling process and the anthropometric techniques used to generate it.

2. Method

2.1. CARAMEX face database

This face database consists of a large sample of facial photographs of individuals from different populations in Mexico. This database was collected between 1993 and 1997 within "La cara del mexicano" (The Mexican face) project using a standardized protocol described in detail previously (Serrano et al., 2000; Farrera et al., 2016), in which photographs were taken at a constant distance of 2 m between the subject and the camera, constant lighting, and the head oriented to the Frankfurt horizontal plane. The entire database was used to generate the composite faces, while randomly selected subsets of faces were used as stimuli for the linguistic/verbal descriptor task.

2.2. Linguistic informants

To build the basic corpus of facial descriptors, 49 students (41 females and 8 males aged from 18 to 25 years old, mean = 21.44, SD = 2.36) provided linguistic/verbal descriptions of a randomly selected subsets of 50 faces. They were recruited through messages posted in the social media of the Psychology Department (i.e., facebook, twitter) and through direct invitations to participate in the study in classrooms at the university.

3. Procedure

3.1. Facial composites

The generation of synthetic faces (portraits) was done in two stages. The first stage involved the syntactic analysis of text files that contained the spatial information of 1) face-shape landmarks and 2) the markers of other face features, of about 1400 frontal images from the CARAMEX collection. In the second stage, clipping areas were set in the original CARAMEX faces to cut and superimpose them on the images of 29 anthropometrically defined face-shape templates to make corrections in terms of scale, position and brightness. On a first approximation, the corrections were made averaging height and width of the faces to be matched with respect to the size of the features, in the case of the positions, and the overall average brightness and contrast of all the facial features on the face-shape templates. All the corrections were performed with the computer vision library OpenCV (see Figure 1).

After all the images were corrected and aligned, we obtained a set of approximately 38,000 synthetic faces from all the possible combinations of facial features within the clipping areas (after discarding some combinations with images that were too dark or that did not scale well).

3.2. Classification test with a convolutional neural network

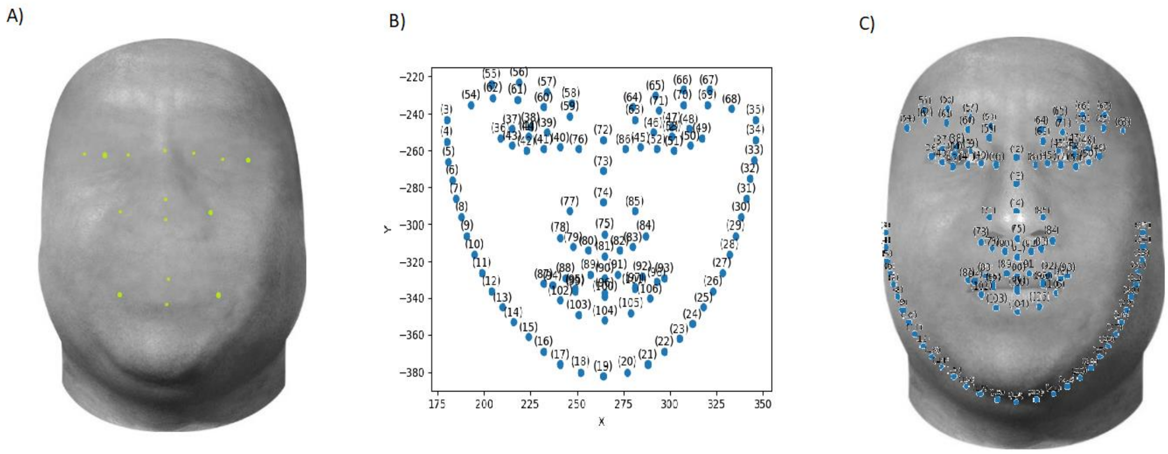
The synthetic faces were used to train and test two convolutional neural networks in a classification task, one supervised and the other unsupervised. For the supervised classification we used the 29 face-shape templates as the criterion.

3.3. Facial linguistic/verbal descriptors

For each participant, fifty faces were selected at random from the CARAMEX database and placed on an internet platform to display them one at a time; the participant saw the face at the center of the screen, and a set of six text boxes on the right in which they could type their personal description of face shape, eyebrows, eyes, nose, mouth, and other distinctive features. Although the description of the set of 50 faces could be done in different sessions according to the participant's available time, they were asked to complete it in one week.

p-ISSN: 2184-2205 e-ISSN: 2184-3414 ISBN: 978-989-35106-0-5 © 2023

Figure 1. Examples of: A) a face-shape template from CARAMEX collection, with landmark detection by OpenCV library; B) landmarks detected by OpenCV on a random face from CARAMEX collection, acting as delimiters for clipping areas; and C) a synthetic face from combination of features from many faces over the face-shaped template.



4. Results

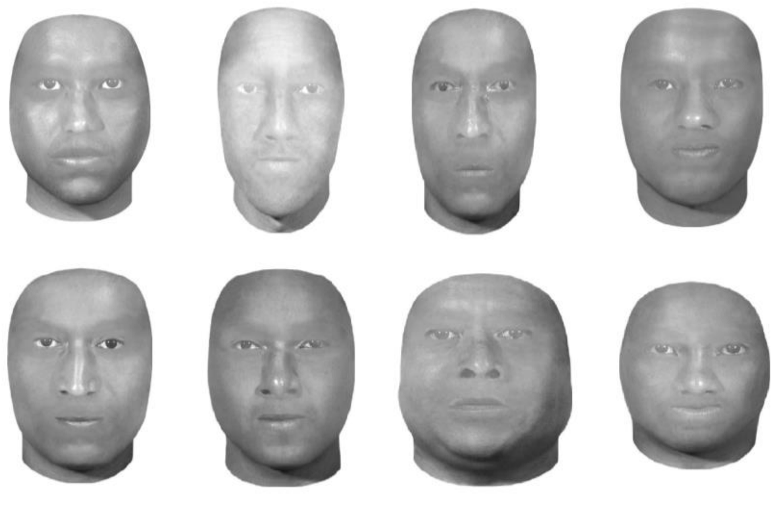
4.1. Synthetic faces database

As mentioned in the methods section, we generated a set of 38,00 facial composites (see figure 2 for some examples).

4.2. Results from the CNN classification task

We tested the accuracy and precision of the two supervised and unsupervised convolutional neural networks using the set of generated images. In the first net, the accuracy obtained was 0.9109 and loss of 0.1691, meanwhile, the unsupervised obtained an accuracy of 0.7794 and loss 2.5819.

Figure 2. A few examples of the synthetic faces generated from a combination of features of CARAMEX faces.

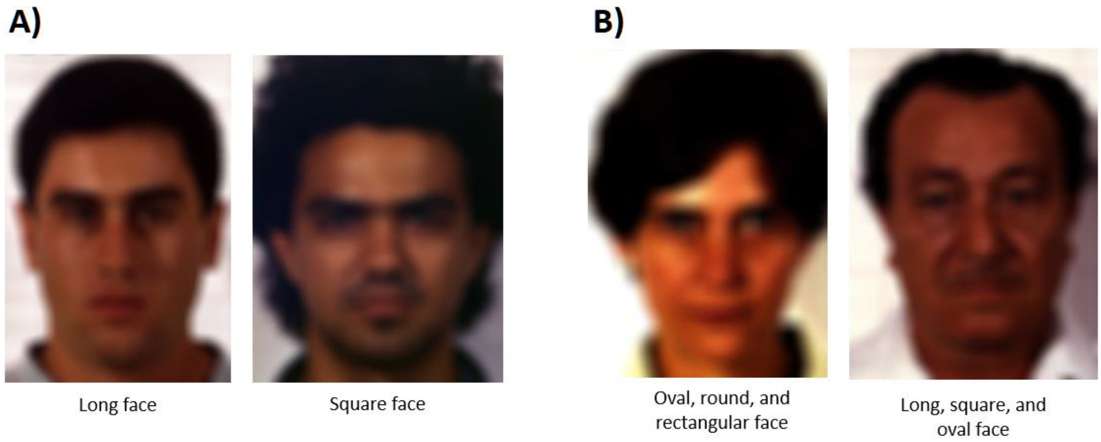


4.3. Corpus of linguistic/verbal facial descriptions

The linguistic/verbal descriptor task resulted in 1331 descriptions. Facial descriptors with a close meaning were considered synonymous (for example, “chica” and “pequeña” or “tupida” and “poblada” in Spanish). These standardized descriptors were analyzed with descriptive statistics.

In this corpus, 457 out of the 1331 descriptions refer to the shape of the face, 443 to the mouth, 1236 to the eyebrows, 437 to the eyes, and 565 to the nose. Round facial shapes (n = 152), medium-sized mouth (n=215), bushy eyebrows (n=498), small eyes (n=134), and broad nose (n=134) were the most frequently observed in this corpus, while trapezoidal facial shapes (n=1), heart-shaped mouths (n=1), unibrows (n=7), slant eyes (n=4) and upturned noses (n=7) were least frequently observed. Next, using only faces described by more than three people (n = 12), we compared the standardized descriptors to the observed facial variation. Despite the small sample size, we can observe that some facial features can be described using a single descriptor (Figure 3A), while others require multiple descriptors (Figure 3B).

Figure 3. Examples of facial descriptions generated for face shape using a single descriptor (A) and multiple descriptors (B). The faces are blurred to protect the privacy of the participants.



5. Conclusions

Despite increasing efforts to improve facial composite construction, forensic face recognition rates remain very low. In this contribution we propose an interdisciplinary and novel approach for the generation of facial composites that, unlike previous approaches using artificial intelligence and/or advanced image processing techniques, introduces both relevant anthropometric variability in the generation of facial composites, and a corpus of linguistic/verbal descriptors to guide a search process.

We showed here the first steps towards this goal: the generation of a database of 38,000 facial composites, and a preliminary linguistic/verbal corpus of facial descriptors obtained through crowdsourcing. Further refinement of the facial composites as well as a larger corpus of linguistic/verbal descriptors will certainly be required.

In future work, we plan to use genetic algorithms to evolve the faces in this database and generate an optimal facial composite in terms of its degree of similarity to the suspect face. Likewise, we also plan to use natural language processing to take advantage of the information in the corpus to guide the evolution through the face space, in order to find an optimized solution that best matches a suspect's face and the victim's or eyewitness' description of said face. With this work we hope to improve the most subjective forensic technique that is routinely carried out in criminal investigations, that is quite susceptible to cognitive biases in both the construction of a facial composite and in the recognition process, with very high social costs (e.g., Cardozo, 2009).

Acknowledgements

We thank DGAPA-UNAM's financial support through the PAPIIT Program for project IT1022422.

References

Baker, M. A., & Reysen, M. B. (2020). The influence of recall instruction type and length on the verbal overshadowing effect. *American Journal of Forensic Psychology, 38*(3), 3-29.

Brown, C., Portch, E., Nelson, L., & Frowd, C. D. (2020). Reevaluating the role of verbalization of faces for composite production: Descriptions of offenders matter!. *Journal of Experimental Psychology: Applied*, *26*(2), 248.

Bruce, V., Ness, H., Hancock, P. J., Newman, C., & Rarity, J. (2002). Four heads are better than one: Combining face composites yields improvements in face likeness. *Journal of Applied Psychology*, *87*(5), 894.

Cardozo, B. N. (2009). Reevaluating lineups: Why witnesses make mistakes and how to reduce the chance of a misidentification. *An Innocence Project Report*.

Farrera, A., García-Velasco, M., & Villanueva, M. (2016). Quantitative assessment of the facial features of a Mexican population dataset. *Forensic Science International*, *262*, 283-e1.

Frowd, C. D. (2021). Forensic facial composites. In *Methods, Measures, and Theories in Eyewitness Identification Tasks* (pp. 34-64). Routledge.

Frowd, C. D., McQuiston-Surrett, D., Anandaciva, S., Ireland, C. G., & Hancock, P. J. (2007). An evaluation of US systems for facial composite production. *Ergonomics*, *50*(12), 1987-1998.

Klum, S., Han, H., Jain, A. K., & Klare, B. (2013, June). Sketch based face recognition: Forensic vs. composite sketches. In *2013 international conference on biometrics (ICB)* (pp. 1-8). IEEE.

Meissner, C. A., & Brigham, J. C. (2001). A meta-analysis of the verbal overshadowing effect in face identification. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *15*(6), 603-616.

Meltzer, M. A., & Bartlett, J. C. (2019). Holistic processing and unitization in face recognition memory. *Journal of Experimental Psychology: General*, *148*(8), 1386–1406

Serrano, C., Villanueva, M., Luy, J., Karl, F. L. (2000). Sistema computarizado de identificación personal con rasgos morfológicos faciales, Antropología Física Latinoamericana 2: 119–134.

Tanaka, J. W., & Simonyi, D. (2016). The “parts and wholes” of face recognition: A review of the literature. *Quarterly Journal of Experimental Psychology*, *69*(10), 1876-1889.

Zahradnikova, B., Duchovicova, S., & Schreiber, P. (2018). Facial composite systems. *Artificial Intelligence Review*, *49*(1), 131-152.