frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 08 October 2010
doi: 10.3389/fpsyg.2010.00159

Interattribute distances do not represent the identity of real-world faces

Vincent Taschereau-Dumouchel[1], Bruno Rossion[2], Philippe G. Schyns[3] and Frédéric Gosselin[1]*

[1] *Département de psychologie, Université de Montréal, Montreal, QC, Canada*
[2] *Unité Cognition et Développement, Institut de Recherches en Sciences Psychologiques et Institut de Neurosciences, Faculté de Psychologie, Université Catholique de Louvain, Louvain-la-Neuve, Belgium*
[3] *Centre for Cognitive Neuroimaging, Department of Psychology, University of Glasgow, Glasgow, UK*

Edited by:
Rufin VanRullen, Centre de Recherche Cerveau et Cognition, Toulouse, France

Reviewed by:
Gyula Kovács, Budapest University of Technology, Hungary
Jesse Husk, McGill University, Canada

Correspondence:
Frédéric Gosselin, Département de psychologie, Université de Montréal, Montreal, Canada.
e-mail: frederic.gosselin@umontreal.ca

According to an influential view, based on studies of development and of the face inversion effect, human face recognition relies mainly on the treatment of the distances among internal facial features. However, there is surprisingly little evidence supporting this claim. Here, we first use a sample of 515 face photographs to estimate the face recognition information available in interattribute distances. We demonstrate that previous studies of interattribute distances generated faces that exaggerated by 376% this information compared to real-world faces. When human observers are required to recognize faces solely on the basis of real-world interattribute distances, they perform poorly across a broad range of viewing distances (equivalent to 2 to more than 16 m in the real-world). In contrast, recognition is almost perfect when observers recognize faces on the basis of real-world information other than interattribute distances such as attribute shapes and skin properties. We conclude that facial cues other than interattribute distances such as attribute shapes and skin properties are the dominant information of face recognition mechanisms.

Keywords: face recognition, interattribute distances

INTRODUCTION

According to an influential view, human face processing rests mainly on *interattribute distances*[1] (e.g., interocular distance, mouth-nose distance; Diamond and Carey, 1986; Carey, 1992; Maurer et al., 2002). After briefly reviewing the origin of this claim, we will examine two points that have perhaps surprisingly been neglected so far: If face processing relies on interattribute distances then, surely, (1) real-world interattribute distances must contain useful information for face recognition; and (2) human observers must be more sensitive to these natural variations than to those of other facial cues.

The origin of the idea that relative distances between features are important for individual face processing can be traced back to the work of Haig (1984), and Diamond and Carey (1986). Haig (1984) moved the different features of a few unfamiliar faces independently by small amounts and measured the just noticeable differences of five observers for all manipulations (e.g., mouth-up, eyes inward) with respect to the original face. He noticed that the sensitivity of human adults to slight alterations in the positions of the features of a set of faces was quite good, at the limit of visual acuity for some alterations (e.g., mouth-up). However, the ranges of these manipulations were arbitrary with respect to normal variations of feature positions in real-life, and there was no assessment of the critical role of such manipulations in actual face identification tasks relative to featural changes.

Based on their developmental studies and their work on visual expertise with non-face objects, Diamond and Carey (1977, 1986) hypothesized that what makes faces special compared to other object categories is the expert ability to distinguish among individuals of the category (i.e., different faces) based on what they so-called "second-order relational properties", namely the idiosyncratic variations of distances between features. However, while these authors claimed that the ability to extract such second-order relational properties would be at the heart of our adult expertise in face recognition (Diamond and Carey, 1986; Carey, 1992), they did not test this hypothesis in any study.

Studies of face inversion have also contributed to the idea that relative distances between attributes are fundamental for face recognition. Faces rotated by 180° in the picture-plane induce important decreases in recognition accuracy and increasing response latencies (e.g., Hochberg and Galper, 1967). This impaired performance is disproportionately larger for faces in contrast to other mono-oriented objects such as houses and airplanes (Yin, 1969; Leder and Carbon, 2006; Robbins and McKone, 2006; for a review see Rossion, 2008, 2009). Thus face inversion has been used as a tool to isolate what is special about upright face processing. It happens that the processing of interattribute distances is more affected by inversion

[1]We intentionally avoided to use the expression "configuration" because it is ambiguous in the face recognition literature: It can refer to either to the relative distances between attributes (e.g., Maurer et al., 2002), or to a way of processing the face ("configural processing", as used by Sergent, 1984; Young et al., 1987) – i.e., as a synonym of "holistic" or as a Gestalt. All face cues, including attribute shapes and skin properties, are "configural" under the latter interpretation. By "interattribute distances", we mean relative distances between facial attributes that can be manipulated independently from the shapes of these attributes (e.g., the center of gravity to center of gravity interocular distance; e.g., Haig, 1984; Sergent, 1984; Hosie et al., 1988; Rhodes et al., 1993; Tanaka and Sengco, 1997; Leder and Bruce, 1998; Freire et al., 2000; Leder and Bruce, 2000; Barton et al., 2001; Leder et al., 2001; Le Grand et al., 2001; Bhatt et al., 2005; Goffaux et al., 2005; Hayden et al., 2007). This excludes, for example, the nasal-corner-to-nasal-corner interocular distance and the temporal-corner-to-temporal-corner interocular distance that cannot be manipulated jointly and independently from attribute size.

than the processing of the local shape or surface-based properties of attributes (Sergent, 1984; Barton et al., 2001; Le Grand et al., 2001; Rhodes et al., 2007; for recent reviews, see Rossion, 2008, 2009).

This last observation has been taken as supporting the view that relative *distances* between facial features are fundamental or most diagnostic for individuating upright faces (e.g., Diamond and Carey, 1977, 1986). However, two critical links are missing in the reasoning. First, there is no direct evidence that interattribute distances are diagnostic for upright face recognition. In fact, there is tentative evidence that interattribute distances might not be the main source of information for face recognition: Exaggerated interattribute distances do not impair recognition much (Caharel et al., 2006); interattribute distances are less useful in similarity judgments than attribute shape (Rhodes, 1988). Second, there is no direct evidence that a difficulty to process interattribute distances is the *cause* of the FIE. In fact, this difficulty can be predicted by Tanaka and Farah (1993) and Farah et al. (1995) proposal that face inversion leads to a loss of the ability to process the face as a gestalt or "holistically" (see Rossion, 2008, 2009, for a discussion). To address the issue of the reliance of face processing on interattribute distances, a first question should be how much do faces vary in interattribute distances in the real-world? Clearly, if there is little objective, real-world interattribute variation, there is little that the visual system could and should do with it.

To address this question, in Experiment 1 we estimated, from a sample of 515 full-frontal real-world Caucasian faces, how much information was objectively available to the human brain in relative interattribute distances for gender discrimination and face identification. We demonstrate that while there is objective interattribute distance information in faces, most previous studies have grossly exaggerated this information when testing it (on average 376%). In Experiment 2, we compared face recognition when interattribute distances are the only information source available (Experiment 2a) and unavailable (Experiment 2b), and we show that performance is much better in the latter case.

EXPERIMENT 1

METHODS

Participants

Three female students (all 19-year-old) from the Université de Montréal received course credits to annotate digital portraits on 20 internal facial feature landmarks (see Experiment 1, Procedure). The first author (22-year-old) annotated faces from previous studies in which the distances between internal features had been altered. Participants had normal or corrected to normal vision.

Stimuli

A total of 515 Caucasian frontal-view real-world portraits presenting a neutral expression (256 females) were used. These faces came from multiple sources: the entire 300-face set of Dupuis-Roy et al. (2009), 146 neutral faces from the Karolinska Directed Emotional Faces, the 16 neutral faces from Schyns and Oliva (1999), the 10 neutral faces from the CAFE set, 6 neutral faces from the Ekman and Friesen (1975) set, and 40 additional neutral faces. We also annotated 86 stimuli used in 14 previous studies in which interattribute distances had been manipulated "within the limit of plausibility" (for a list, see **Figure 2**).

Apparatus

The annotations were made on a Macintosh G5 computer running functions written for Matlab (available at http://mapageweb.umontreal.ca/gosselif/alignTools/) using functions from the Psychtoolbox (Brainard, 1997; Pelli, 1997). Stimuli were presented on a HP p1230 monitor at a resolution of 1920×1200 pixels with a 100-Hz refresh rate. The monitor luminance ranged from 1.30 to 80.9 cd/m^2.

Procedure

Participants were asked to place 20 points on specific landmarks of internal facial features with a computer mouse, one face at a time (see blue crosses in the leftmost column of **Figure 1**). These landmarks were chosen because they are easy to locate and allow a proper segmentation of the features (Okada et al., 1999). (If we had to do it again, however, we would use four landmarks instead of two for the eyebrows.) We increased the size of stimuli to match computer monitor resolution to ease the task of participants. Every participant annotated each of the 515 portraits in random order allowing us to estimate inter-subject annotation error.

RESULTS

We reduced each set of 20 annotations to 6 feature positions by averaging the *xy*-coordinates of the annotations placed on landmarks belonging to the same facial feature (see green dots in the leftmost column of **Figure 1**) to disentangle attribute position from attribute shape. Indeed, to manipulate interattribute distances independently from attribute shape, whole attributes are typically cropped – including, in our case, all the pixels annotated by our observers on each of these attributes – and translated (e.g., Maurer et al., 2002).

This 20-to-6 reduction also maximizes signal-to-noise ratio of attribute position. Assuming that annotation error is the same for the *x*- and *y*-dimensions and for all features (and systematic error aside), the signal-to-noise ratio of the measurements is estimated at 8.27 per annotation (i.e., $(\sigma^2_{\text{total}} - \sigma^2_{\text{inter-participantl}})/\sigma^2_{\text{inter-participant}} = 8.27$, with $\sigma^2_{\text{total}} = 8.80$ and $\sigma^2_{\text{inter-participant}} = 0.95$ pixels per annotation for a mean interocular distance of 100 pixels). For all attributes except the eyebrows, four annotations were averaged, and thus signal-to-noise ratio of attribute position is twice that for individual annotations (i.e., 16.54); for the eyebrows, two annotations were averaged, and thus, signal-to-noise ratio was $\sqrt{2}$ that for individual annotations (i.e., 11.70). In sum, the signal-to-noise ratio of attribute position was high.

To estimate absolute interattribute distances, the brain would have to estimate absolute depth precisely; such absolute depth estimates are only possible at really close range, which is atypical of face identification distances. Thus it is usually assumed that only *relative* interattribute distances are available to the brain for face identification (Rhodes, 1988).

We "relativized" interattribute distances by translating, rotating, and scaling the feature positions of each face to minimize the mean square of the difference between them and the average feature positions across faces (rotated so that the *y*-axis was the main facial axis; see the rightmost column of **Figure 1**; Ullman, 1989). Technically, this is a *linear conformal transformation*; it preserves relative interattribute distances (e.g., Gonzalez et al., 2009). This procedure is implemented in the companion Matlab functions (http://mapageweb.umontreal.ca/gosselif/alignTools/).

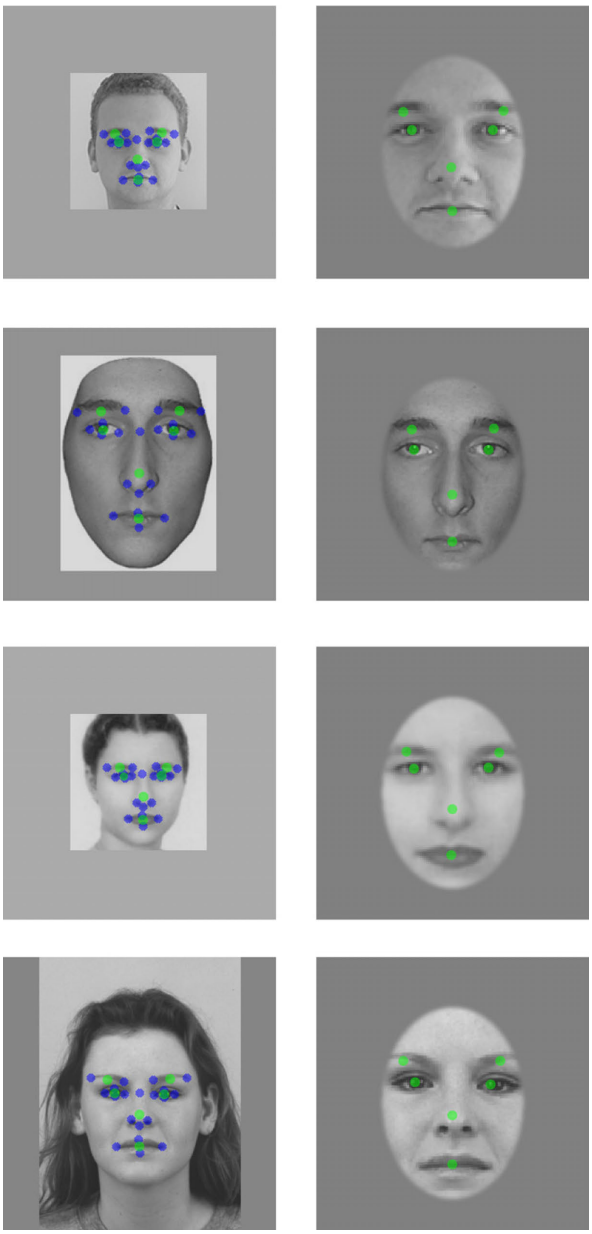

FIGURE 1 | Annotation and alignement procedure. Leftmost column: Sample faces annotated in Experiment 1. The 20 blue crosses show, for these faces, the average annotations across participants. These 20 annotations were reduced to six attribute positions – green dots – by averaging the coordinates of the annotations belonging to every attribute. Rightmost column: We translated, rotated, and scaled the attribute positions of each face to minimize the mean square of the difference between them and the average attribute positions across faces. The residual differences between aligned attribute positions – green dots – is the interattribute variance in the real-world.

The resulting interattribute distances are proportional to the ones obtained by dividing the interattribute distances of each face by its mean interattribute distance. However, our alignment procedure provides an intuitive way of visualizing the *variance* of interattribute distances. The green dots in **Figure 2** represent the distributions of the aligned feature positions of real-world faces.

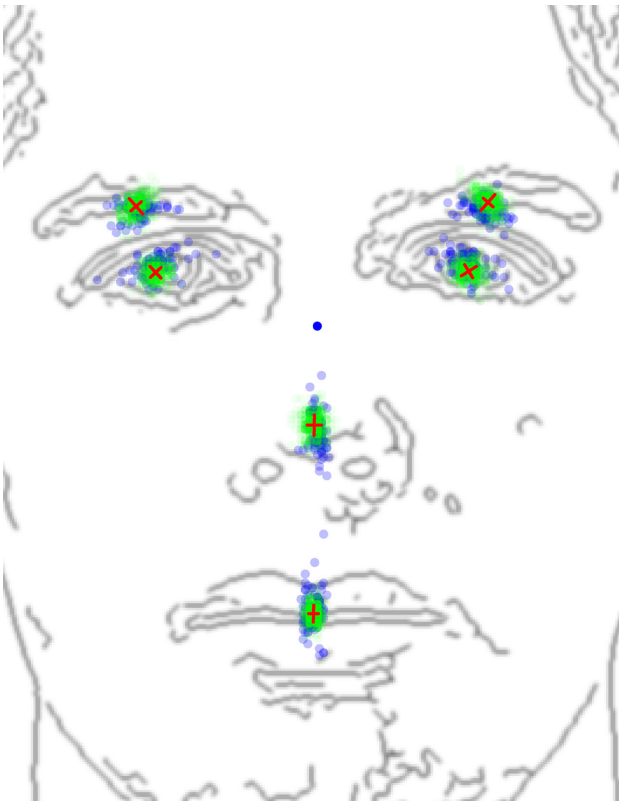

FIGURE 2 | Distribution of interattribute distances. Distribution of post-alignment attribute positions (green dots) of the 515 annotated faces, with standard-deviation-length eigenvectors (red segments) centered on the distributions, and overlaid to the contours of a face to facilitate the interpretation. The blue dots are the attribute positions of stimuli of 14 previous studies that used distance manipulations (Haig, 1984; Sergent, 1984; Hosie et al., 1988; Rhodes et al., 1993; Tanaka and Sengco, 1997; Leder and Bruce, 1998, 2000; Freire et al., 2000; Barton et al., 2001; Leder et al., 2001; Le Grand et al., 2001; Bhatt et al., 2005; Goffaux et al., 2005; Hayden et al., 2007).

The variance of each distribution reflects the contribution of the corresponding attribute to the overall interattribute distance variance in the real-world. (See Section "How to Create Realistic Interattribute Distances" in Appendix for a description of the *covariance* between the aligned attributes.) Red lines represent one standard deviation of the aligned positions along their first and second components of the principal component analysis (PCA). As can be seen at a glance, the pairs of red lines on the eyes and eyebrows are of similar lengths, which means that the variance in the positions of these features is roughly the same at all orientations. However, the pairs of red lines on the nose and mouth are clearly of different lengths – for these features, the variance is mainly organized along the main facial axis.

We also plotted the aligned attribute positions of 86 artificial stimuli drawn from 14 studies that explicitly manipulated interattribute distances (blue dots). The mean distances between these artificial dots and the natural dots, expressed in standard deviations of the natural dots, is 0.831 (SD = 1.497). On average, the eyes

along the main facial axis diverged most (on the right of the image: mean = 2.147, SD = 1.244; on the left of the image: mean = 1.633, SD = 1.240). More than 73% of the experimental faces had at least one attribute falling more than two standard deviations away from the mean of at least one axis of the real-world faces (23% of the eyes on both axes, and 26% of the noses and 29% of the mouths on the *y*-axis). Thus, in most of these 14 studies, artificial interattribute distances were exaggerated compared to natural variations. What was the impact of this exaggeration on the information available for face recognition?

To answer this question, we performed two virtual experiments (see Section "Virtual Experiments" in Appendix for details). In the first one, we repeatedly trained a model at identifying, solely on the basis of interattribute distances, one randomly selected natural face from 50% of the natural faces, also randomly selected, and tested the model on the remaining natural faces. Similarly, in the second virtual experiment, we repeatedly trained a model at identifying one randomly selected *artificial* face from 50% of the *natural* faces, also randomly selected, and tested the model on the remaining *natural* faces. In each case, we found how much noise was necessary for the models to perform with a fixed sensitivity ($A' = 0.75$). The models trained to identify the artificial faces required about 3.76 times more noise ($\sigma^2 = 155.40$ pixels for a mean interocular distance of 100 pixels) than the ones trained to identify the real-world faces ($\sigma^2 = 41.28$ pixels for a mean interocular distance of 100 pixels). This implies that the interattribute distances of the artificial faces convey about 376% more information for identification than in real-world faces.

In sum, there is information in interattribute distances for processing in real-world faces. However, not nearly as much as the majority of past studies have assumed.

EXPERIMENT 2

In Experiment 2a, we asked whether human observers can use this real-world interattribute distance information to resolve a matching-to-sample (ABX) task when interattribute distance is the only information available. And, in Experiment 2b, we asked the complementary question: Can human observers use real-world cues other than interattribute distances such as attribute shapes and skin properties to resolve an ABX?

METHODS

Participants

Sixteen observers (eight females and eight males; aged between 19 and 29 years of age; mean = 22.8 years; SD = 2.5 years) participated in Experiment 2a; and 10 different observers (five females and five males; aged between 21 and 31 years of age; mean = 23.9 years; SD = 3.28 years) participated in Experiment 2b. All observers had normal or corrected to normal vision.

Stimuli

We created 2,350 pairs of stimuli for each experiment. Base faces were those annotated in Experiment 1. First, we translated, rotated, and scaled all these face images to minimize the mean square of the difference between their feature positions (their 20 annotations distilled to six attributes' centers of gravity) and the average feature positions across faces rescaled to an interocular distance of 50 pixels (or 1.4 cm). Technically, we performed *linear conformal transformations*, which preserve relative interattribute distances (e.g., Gonzalez et al., 2009). To create one stimulus pair in Experiment 2a, we randomly selected three faces of the same gender from the bank of 515 faces. We chose faces of the same gender to eliminate the possible gender discrimination confound. We cut out the six attributes of one of these faces – the *feature face* – we displaced them to the locations of the attributes of one of the two remaining faces – the first *distance face* – and we filled in the holes to create the first stimulus; and then we displaced the six attributes of the *feature face* to the locations of the attributes of the third face – the second *distance face* – and we filled in the holes to create the second stimulus. This procedure ensures that face stimuli from a pair have identical internal features and only differ on the distances between these features. More specifically, feature masks were best-fitted to the annotations of every internal features of one of the *feature face* – using *affine transformations* (e.g., Gonzalez et al., 2009). The pixels covered by the features masks were then translated to the feature positions of the two *distance faces* – producing a pair of face stimuli (see **Figure 3** – see Section "How to Create Realistic Interattribute Distances" in Appendix for an alternative method for creating realistic interattribute distances). Pixels falling outside the feature models were inferred from the *feature face* using bicubic interpolation (e.g., Keys, 1981). This procedure is implemented in the companion Matlab functions.

In Experiment 2b we also randomly selected three faces of the same gender from the database. This time, however, we best-fitted feature models to the landmarks of the internal features of two of these faces – the *feature faces* – and the features were translated according to the feature positions of the third face – the *distance face*. The pixels falling outside the feature models were interpolated from the appropriate *feature face*. This procedure ensures that faces from a stimulus pair have identical interattribute distances but differ in cues other than interattribute distances, such as attribute shapes and skin properties.

All face stimuli were shown in grayscale, with equal luminance mean and variance, through a gray mask punctured by an elliptic aperture with a smooth edge (convolved with a Gaussian kernel with a standard deviation equal to 2 pixels) and with a horizontal diameter of 128 pixels and a vertical diameter of 186 pixels. This only revealed the inner facial features and their distances (for examples, see **Figure 4**).

Apparatus

Experiment 2 was performed on a Macintosh G5 running a computer script written for the Matlab environment using functions of the Psychtoolbox (Brainard, 1997; Pelli, 1997). Stimuli were presented on a HP p1230 monitor at a resolution of 1920×1200 pixels at a refresh rate of 100 Hz. The monitor luminance ranged from 1.30 to 80.9 cd/m^2.

Procedure

Participants completed 120 trials of their ABX task (the sequence of events in a trial is given in **Figure 4**) at each of five viewing distances in a randomized block design to equate the effect of learning. Viewing distances were equivalent to real-world viewing distances of 2, 3.4, 5.78, 9.82, and 16.7 m (which corresponds, respectively, to average interocular widths of 1.79°, 1.05°, 0.62°, 0.37°, and 0.21° of visual angle). This represent a broad range of viewing distances

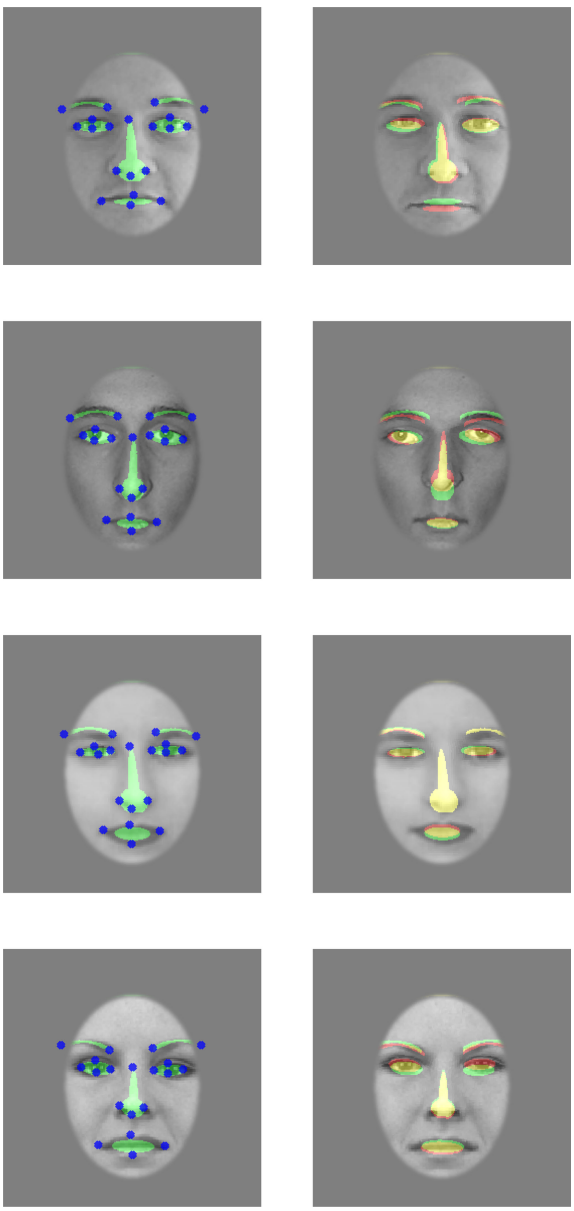

FIGURE 3 | Creation of face stimuli using real-world interattribute distances. Leftmost column: In Experiment 2, feature masks – shown in translucid green – were bestfitted to the aligned annotations – represented by blue crosses. Rightmost column: In Experiment 2a, these feature masks were displaced according to the feature positions of another face of the same gender. Translucid green areas reproduce the feature masks of the leftmost column; translucid red areas represent the same feature masks after displacement; and translucid yellow areas represent the overlap between these two sets of feature masks.

from which faces can be readily recognized; one reason for including a variety of viewing distances was to test whether the use of interattribute distances is indeed invariant to viewing distances (Rhodes, 1988). We used the interocular width average of 6.2 cm (mean for males = 6.3 cm; and mean for females = 6.1 cm) reported by Farkas (1981) to determine the equivalent real-world distances.

On each trial, one stimulus from a pair (see Experiment 2, Stimuli) was randomly selected as the target. This target was then presented for 800 ms immediately followed by a blank presented for 200 ms immediately followed by the pair of stimuli presented side-by-side in a random order. The pair of stimuli remained on the screen while participants were asked to choose which face – on the left or on the right – was the target. No feedback was provided to the participants between trials.

RESULTS

We submitted the results to a 2 × (5) mixed design ANOVA using *viewing distances* (2, 3.4, 5.78, 9.82, and 16.7 m) as a within-subjects factor and *group* (different vs. same interattribute distances) as a between-subjects factor. Contrasts of the between-subjects factor revealed a significant difference of accuracy between the two groups at all five viewing distances [all $F(1,24) > 100$, $p < 0.00001$, $\eta^2 > 0.80$, $p_{rep} \approx 1$]. Observers who had to rely solely on interattribute distances performed significantly lower than observers who had to use features at each of the five viewing distances (see **Figure 5**). There was also a significant interaction between viewing distances and groups [$F(2.4; 57.7) = 4.89$, $p = 0.007$, $\eta^2 = 0.17$, $p_{rep} = 0.96$].

To test the effect of distance of presentation within each group, the data was separated and one-way ANOVAs were carried on each group independently. The within-subjects analysis revealed no differences of accuracy between any viewing points in the task where the interattribute distances were kept constant [$F(1.9,17.2) = 1.60$, *ns*]. The same analysis revealed a significant difference between response accuracy as a function of distance in the group where the interattribute distances were different [$F(2.3,35) = 10.51$, $p = 0.0001$, $\eta^2 = 41$, $p_{rep} = 0.98$]. A polynomial contrast revealed a significant linear relationship between response accuracy and distance of presentation when interattribute distances is the sole information available to perform the discrimination [$F(1, 15) = 24.41$, $p = 0.0001$, $\eta^2 = 0.62$, $p_{rep} = 0.98$]. The group averages in this task indicated a decreasing accuracy with increasing distances (*nearest*: mean = 64.74%, SD = 8.6; *furthest*: mean = 55.1%, SD = 5.77). **Figure 5** displays the mean proportions and standard errors of correct responses as a function of viewing distances.

A 2 × (5) mixed ANOVA with *viewing distances* (2, 3.4, 5.78, 9.82, and 16.7 m) as a within-subjects factor and *groups* (different vs. same interattribute distances) as a between-subjects factor revealed a main effect of groups on response time [$F(1,24) = 23.81$, $p < 0.00001$, $\eta^2 = 0.50$, $p_{rep} = 0.99$]. Same interattribute distances (mean = 0.995 s, SD = 0.22) elicited significantly faster reaction times than different interattribute distances (mean = 1.94 s, SD = 0.73).

GENERAL DISCUSSION

In Experiment 1, we asked whether relative distances between real-world internal facial features contain enough information for face categorizations (identity and gender). We carried out a series of simulations on these faces to assess the information available in their residual interattribute distances. We found that real-world interattribute distances did in fact contain information useful to resolve face identification.

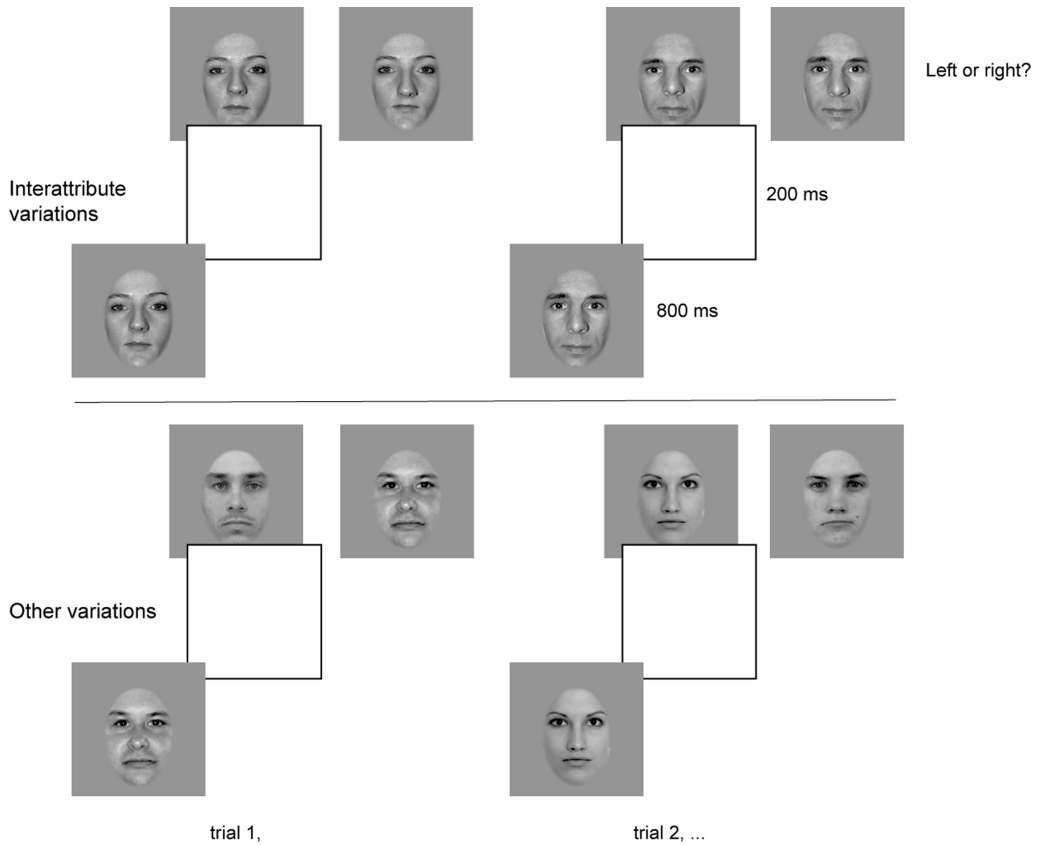

FIGURE 4 | Sequence of events in two sample trials of our experiments. Top: In Experiment 2a, we asked whether human observers can use this real-world interattribute distance information, at different viewing distances, to resolve a matching-to-sample (ABX) task when interattribute distance is the only information available. Bottom: In Experiment 2b, we asked the complementary question: Can human observers use real-world cues other than interattribute distances such as attribute shapes and skin properties, at different viewing distances, to resolve an ABX?

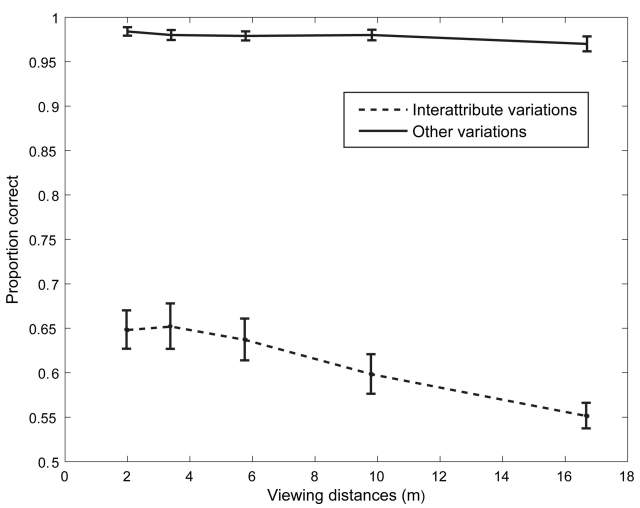

FIGURE 5 | Mean proportion of correct face recognition in function of distance. Error bars represent one standard error. The dashed line represents performance when real-world interattribute distance is the only information available (Experiment 2a); and the solid line represents performance when only real-world cues other than interattribute distances such as attribute shapes and skin properties are available.

In Experiment 2a, we examined whether human observers could use real-world interattribute distance information to resolve a matching-to-sample (ABX) task when this is the only information available. In Experiment 2b – the exact reciprocal of Experiment 2a – we asked if human observers could use information other than interattribute distances, such as attribute shapes and skin reflectance properties, to resolve an ABX task. Results of the Experiment 2a indicated that human observers perform poorly when required to recognize faces solely on the basis of real-world interattribute distances at all tested viewing distances (equivalent to 2 to more than 16 m in the real-world, a broad range of viewing distances from which faces can be readily recognized) (best accuracy = 65% correct); whereas results of Experiment 2b showed that they perform close to perfection when required to recognize faces on the basis of real-world information other than interattribute distances such as attribute shapes and skin properties (e.g., O'Toole et al., 1999) at all tested viewing distances. Moreover, the performance of human observers decreased linearly with increasing viewing distances when required to recognize faces solely on the basis of real-world interattribute distances. If interattribute distances appealed to researchers as a face representation code it is in part because they are invariant to viewing distances (e.g., Rhodes, 1988). Human observers seem incapable to take advantage of this property of interattribute distances.

One reason that may explain why the majority of researchers overestimated the importance of interattribute distances for face recognition is the use of grotesque face stimuli. We have computed that on average face stimuli in the interattribute distance literature convey 376% more information for identification than real-world faces. These artificial stimuli were created by using various computer tools to crop and to move the internal facial features, for example, increasing or reducing interocular distance and/or mouth/nose distance (e.g., Rhodes et al., 1993, 2007; Freire et al., 2000; Leder and Bruce, 2000; Barton et al., 2001; Leder et al., 2001; Le Grand et al., 2001; Leder and Carbon, 2006; Goffaux and Rossion, 2007; Goffaux, 2008), and to fill in the hole(s) left behind. This kind of transformation does not necessarily respect the range of real-world interattribute distances. In fact, in most studies of face inversion, variations of distances between features were intentionally stretched to the limit of "plausibility" to obtain a reasonably good performance at upright orientation (e.g., Freire et al., 2000; Barton et al., 2001; Goffaux and Rossion, 2007; Rhodes et al., 2007). In an attempt to create more realistic face stimuli, a subset of these researchers (e.g., Le Grand et al., 2001; Mondloch et al., 2002, 2003; Hayden et al., 2007) altered the interocular or the nose-to-mouth distance within a reasonable number of standard deviations from the mean of the anthropometric norms of Farkas (1981). However, this effort is insufficient because the Farkas statistics are contaminated by an undesirable source of variance – the absolute size of faces which is unavailable to the brain of observers as we explained above and thus irrelevant to face recognition – and they do not contain covariance information – e.g., the fact that the eyebrows have a tendency to follow the eyes. Rather, in our experiments, we have sampled relative interattribute distances from real-world distributions. Alternatively, the method presented in Section "How to Create Realistic Interattribute Distances" in Appendix of this article can be used.

The results of Experiment 2a are all the more remarkable that they provide an *upper-bound* on the usefulness of interattribute distances for real-world face recognition. Our ABX task, which requires the identification of one recently viewed face among two face stimuli, is much easier than real-life face identification, which requires typically the comparison of hundreds of memorized faces with one face stimulus. Furthermore, no noise was added to the interattribute distances of our stimuli; real-life interattribute distances are contaminated by several sources of noise – facial movements, foreshortening, shadows, and so on. Finally, the interattribute distance information of our stimuli slightly overestimated real-life interattribute distance information because of unavoidable annotation errors. In conclusion, facial cues other than interattribute distances such as attribute shapes and skin properties are the dominant information of face recognition mechanisms in the real-world.

Our results fall short of accounting for the poor performance with interattribute distances. It could be that there is less interattribute distance information available to resolve the task or that observers are inept at using interattribute distance information. One approach to compare performance in both conditions fairly would be to measure efficiencies (e.g., Tjan et al., 1995), which are task-invariant indices of performance.

REFERENCES

Barton, J. J., Keenan, J. P., and Bass, T. (2001). Discrimination of spatial relations and features in faces: effects of inversion and viewing duration. *Br. J. Psychol.* 92, 527–549.

Bhatt, R. S., Bertin, E., Hayden, A., and Reed, A. (2005). Face processing in infancy: developmental changes in the use of different kinds of relational information. *Child Dev.* 76, 169–181.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vision* 10, 433–436.

Caharel, S., Fiori, N., Bernard, C., Lalonde, B., and Rebaï, M. (2006). The effect of inversion and eye displacements of familiar and unknown faces on early and late-stage ERPs. *Int. J. Psychophysiol.* 62, 141–151.

Carey, S. (1992). Becoming a face expert. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 335, 95–103.

Diamond, R., and Carey, S. (1977). Developmental changes in the representation of faces. *J. Exp. Child. Psychol.* 23, 1–22.

Diamond, R., and Carey, S. (1986). Why faces are and are not special: An effect of expertise. *J. Exp. Psychol. Gen.* 115, 107–117.

Dupuis-Roy, N., Fortin, I., Fiset, D., and Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *J. Vision* 9, 1–8.

Ekman, P., and Friesen, W. V. (1975). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues.* Englewood Cliffs, NJ: Prentice-Hall.

Farah, M. J., Tanaka, J. W., and Drain, H. M. (1995). What causes the face inversion effect? *J. Exp. Psychol. Hum. Percept. Perform.* 21, 628–634.

Farkas, L. G. (1981). *Anthropometry of the Head and the Face.* New York: Elsevier.

Freire, A., Lee, K., and Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: direct evidence. *Perception* 29, 159–170.

Goffaux, V. (2008). The horizontal and vertical relations in upright faces are transmitted by different spatial frequency ranges. *Acta Psychol.* 128, 119–126.

Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., and Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception* 34, 77–86.

Goffaux, V., and Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 995–1002.

Gonzalez, R. C., Woods, R. E., and Eddins, S. L. (2009). *Digital Image Processing Using MATLAB*, Second Edition. Gatesmark Publishing.

Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception* 13, 502–512.

Hayden, A., Bhatt, R. S., Reed, A., Corbly, C. R., and Joseph, J. E. (2007). The development of expert face processing: are infants sensitive to normal differences in second-order relational information? *J. Exp. Child. Psychol.* 97, 85–98.

Hochberg, J., and Galper, R. (1967). Recognition of faces: I. An exploratory study. *Psychon. Sci.* 9, 619–620.

Hosie, J. A., Ellis, H. D., and Haig, N. D. (1988). The effect of feature displacement on perception of well known faces. *Perception* 17, 461–474.

Keys, R. (1981). Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust. Speech Signal Process.* 29, 1153.

Leder, H., and Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Q. J. Exp. Psychol. A* 51, 449–473.

Leder, H., and Bruce, V. (2000). When inverted faces are recognized: the role of configural information in face recognition. *Q. J. Exp. Psychol. A* 53, 513–536.

Leder, H., and Carbon, C. C. (2006). Face-specific configural processing of relational information. *Br. J. Psychol.* 97, 19–29.

Leder, L., Candrian, G., Huber, O., and Bruce, V. (2001). Configural information in the context of upright and inverted faces. *Perception* 30, 73–83.

Le Grand, R., Mondloch, C. J., Maurer, D., and Brent, H. P. (2001). Early visual experience and face processing. *Nature* 410, 890 (Correction: *Nature* 412, 786).

Maurer, D., Le Grand, R., and Mondloch, C. J. (2002). The many faces of configural processing. *Trends Cogn. Sci.* 6, 255–260.

Mondloch, C. J., Geldart, S., Maurer, D., and Le Grand, R. (2003). Developmental changes in face processing skills. *J. Exp. Child. Psychol.* 86, 67–84.

Mondloch, C. J., Le Grand, R., and Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception* 31, 553–566.

Okada, K., von der Malsburg, C., and Akamatsu, S. (1999). "A pose-invariant face recognition system using linear PCMAP model," in *Proceedings of IEICE Workshop of Human Information Processing (HIP99-48)*, November 1999, Okinawa, 7–12

O'Toole, A. J., Vetter, T., and Blanz, V. (1999). Two-dimensional reflectance and three-dimensional shape contributions to recognition of faces across viewpoint. *Vision Res.* 39, 3145–3155.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics:

transforming numbers into movies. *Spat. Vision* 10, 437–442.
Rhodes, G. (1988). Looking at faces – 1st order and 2nd order features as determinants of facial appearance. *Perception* 17, 43–63.
Rhodes, G., Brake, S., and Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition* 47, 25–57.
Rhodes, G., Hayward, W. G., and Winkler, C. (2007). Expert face coding: configural and component coding of own-race and other-race faces. *Psychon. Bull. Rev.* 13, 499–505.
Robbins, R., and McKone, E. (2006). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition* 103, 34–79.
Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychol.* 128, 274–289.
Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: the perceptual field hypothesis. *Acta Psychol.* 132, 300–312.
Schyns, P. G., and Oliva, A. (1999). Dr Angry and Mr Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition* 69, 243–265.
Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *Br. J. Psychol.* 75, 221–242.
Tanaka, J. W., and Farah, M. J. (1993). Parts and wholes in face recognition. *Q. J. Exp. Psychol. A* 46, 225–245.
Tanaka, J. W., and Sengco, J. A. (1997). Features and their configuration in face recognition. *Mem. Cogn.* 25, 583–592.
Tjan, B. S., Braje, W. L., Legge, G. E., and Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Res.* 35, 3053–3069.
Ullman, S. (1989). Aligning pictorial descriptions: an approach to object recognition. *Cognition* 32, 193–254.
Yin, R. K. (1969). Looking at upside-down faces. *J. Exp. Psychol.* 81, 141–145.
Young, A. M., Hellawell, D., and Hay, D. C. (1987). Configural information in face perception. *Perception* 10, 747–759.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 June 2010; accepted: 09 September 2010; published online: 08 October 2010.
Citation: Taschereau-Dumouchel V, Rossion B, Schyns PG and Gosselin F (2010) Interattribute distances do not represent the identity of real-world faces. Front. Psychology ***1****:159. doi: 10.3389/fpsyg.2010.00159*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology

APPENDIX

VIRTUAL EXPERIMENTS

In the first virtual experiment, we trained, by matrix pseudoinverse, a minimum squared-error linear classifier at identifying one randomly selected natural face from 50% of the remaining natural faces, also randomly selected:

$$\mathbf{b} = (\mathbf{X}_1{}^{\mathrm{T}}\mathbf{X}_1)^{-1}\mathbf{X}_1{}^{\mathrm{T}}\mathbf{y}_1,$$

with $\mathbf{b}$, the regression coefficients of the classifier. $\mathbf{X}_1$ is the set of training vectors:

$$\begin{bmatrix} \mathrm{x}_{\mathrm{target}} & 1 \\ \ldots & 1 \\ \mathrm{x}_{\mathrm{target}} & 1 \\ \mathrm{x}_{\mathrm{distracter1}} & 1 \\ \ldots & 1 \\ \mathrm{x}_{\mathrm{distracter257}} & 1 \end{bmatrix} + \begin{bmatrix} N(0,\sigma) & \ldots & N(0,\sigma) & 0 \\ \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & 0 \\ N(0,\sigma) & \ldots & N(0,\sigma) & 0 \end{bmatrix};$$

with $\mathbf{x}_i$, a training vector corresponding to the set of aligned attribute coordinates of face i: $[x_{\text{left eyebrow}}\ x_{\text{right eye brow}}\ x_{\text{left eye}}\ x_{\text{right eye}}\ x_{\text{nose}}\ x_{\text{mouth}}\ y_{\text{left eyebrow}}\ y_{\text{right eye brow}}\ y_{\text{left eye}}\ y_{\text{right eye}}\ y_{\text{nose}}\ y_{\text{mouth}}]$, where y and x are the coordinates along the main facial axis (y-axis) and the axis orthogonal to it (x-axis); and with $N(0,\sigma)$, a Gaussian random variable of mean 0 and of standard deviation σ. And $\mathbf{y}_1$ is the category membership of the training vectors:

$$\begin{bmatrix} 1 \\ \ldots \\ 1 \\ -1 \\ \ldots \\ -1 \end{bmatrix}.$$

We tested this model on the remaining natural faces:

$$\mathbf{y}_2{}' = f(\mathbf{b}\mathbf{X}_2),$$

with f, a Heaviside step function – if the value exceeds c the function outputs 1 and otherwise it outputs −1. $\mathbf{X}_2$ is the set of test vectors:

$$\begin{bmatrix} \mathrm{x}_{\mathrm{target}} & 1 \\ \ldots & 1 \\ \mathrm{x}_{\mathrm{target}} & 1 \\ \mathrm{x}_{\mathrm{distracter258}} & 1 \\ \ldots & 1 \\ \mathrm{x}_{\mathrm{distracter515}} & 1 \end{bmatrix} + \begin{bmatrix} N(0,\sigma) & \ldots & N(0,\sigma) & 0 \\ \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & 0 \\ N(0,\sigma) & \ldots & N(0,\sigma) & 0 \end{bmatrix}.$$

We computed the hit ($\mathrm{y}_2{}'$ = target and y_2 = target) and false alarm ($\mathrm{y}_2{}'$ = target and y_2 = distracter) rates in function c to obtain a receiver operating characteristic (ROC) curve. Our measure of sensitivity was the area under the ROC curve (A'). We repeated this procedure 20,000 times: We used 20 levels of noise ($\sigma = 1, 2, 3, \ldots, 20$ pixels for a mean interocular distance of 100 pixels); and, for each level of noise, we did 1,000 repetitions in order to obtain a stable estimate of A'. We bestfitted a linearly transformed power function to these data (all $R^2 > 0.99$) and interpolated the quantity of noise required for the classifier to perform with an $A' = 0.75$.

Similarly, in the second virtual experiment, we trained a minimum squared-error linear classifier at identifying one randomly selected artificial face from 50% of the natural faces, also randomly selected, and tested the classifier on the remaining natural faces. Again, we repeated this procedure 20,000 times (1,000 repetitions × 20 levels of noise). Finally, we found the quantity of noise required for the classifier to perform with an $A' = 0.75$.

The classifiers trained to identify the artificial faces required about 3.76 times more noise ($\sigma^2 = 155.40$ pixels for a mean interocular distance of 100 pixels) than the ones trained to identify the real-world faces ($\sigma^2 = 41.28$ pixels for a mean interocular distance of 100 pixels). In other words, the interattribute distances of the artificial faces convey about 376% more information for identification than in real-world faces.

HOW TO CREATE REALISTIC INTERATTRIBUTE DISTANCES

The most straightforward method – the one we opted for in Experiment 2 – consists in sampling interattribute distances from a real-world distribution. This solution has the advantage of preserving all interattribute distance information; but it has the disadvantage of being clumsy. In this appendix, we sketch an alternative method, which is a good compromise.

We can simulate the variance and covariance of the interattribute distances of our aligned female face set with μ_{female} the mean of the xy-coordinates of those features:

$$\begin{bmatrix} -39.10 \\ -39.67 \\ -16.40 \\ -16.50 \\ 32.96 \\ 91.56 \\ -56.86 \\ 56.84 \\ -50.23 \\ 50.37 \\ 0.32 \\ -0.50 \end{bmatrix}$$

(y-coordinates of left eyebrow, right eyebrow, left eye, right eye, nose, and mouth, followed by the x-coordinates of the same; the upper left quadrant being negative for both x- and y-coordinates) and $\boldsymbol{K}_{\text{female}}$ their covariance matrix:

$$\begin{bmatrix}
5.67 & 2.42 & -1.96 & -2.15 & -2.01 & -2.00 & -2.76 & 3.45 & -0.85 & 2.39 & -0.76 & -1.19 \\
2.42 & 6.37 & -2.96 & -2.74 & -1.78 & -1.25 & -3.68 & 2.85 & -1.61 & -0.14 & 1.38 & 1.30 \\
-1.96 & -2.96 & 5.13 & 3.87 & -2.76 & -1.21 & 0.35 & 0.33 & -0.77 & 1.23 & -0.74 & -0.53 \\
-2.15 & -2.74 & 3.87 & 5.79 & -3.16 & -1.31 & -0.40 & -0.49 & -1.33 & 1.48 & 0.39 & 0.24 \\
-2.01 & -1.78 & -2.76 & -3.16 & 10.77 & -1.03 & 2.87 & -2.26 & 1.49 & -1.93 & -0.36 & 0.08 \\
-2.00 & -1.25 & -1.21 & -1.31 & -1.03 & 6.90 & 3.58 & -4.04 & 3.05 & -2.97 & 0.16 & 0.18 \\
-2.76 & -3.68 & 0.35 & -0.40 & 2.87 & 3.58 & 6.68 & -5.11 & -0.11 & -0.45 & -0.83 & -0.25 \\
3.45 & 2.85 & 0.33 & -0.49 & -2.26 & -4.04 & -5.11 & 7.22 & -0.84 & -0.89 & 0.06 & -0.27 \\
-0.85 & -1.61 & -0.77 & -1.33 & 1.49 & 3.05 & -0.11 & -0.84 & 5.42 & -2.79 & -1.26 & -0.25 \\
2.39 & -0.14 & 1.23 & 1.48 & -1.93 & -2.97 & -0.45 & -0.89 & -2.79 & 6.72 & -1.74 & -0.61 \\
-0.76 & 1.38 & -0.74 & 0.39 & -0.36 & 0.16 & -0.83 & 0.06 & -1.26 & -1.74 & 3.89 & -0.27 \\
-1.19 & 1.30 & -0.53 & 0.24 & 0.08 & 0.18 & -0.25 & -0.27 & -0.25 & -0.61 & -0.27 & 1.73
\end{bmatrix}$$

via the following transformation of a Gaussian noise vector **w**:

$$\mathbf{x}_{\text{female}} = E_{\text{female}}\Lambda_{\text{female}}^{1/2}\mathbf{w} + \mu_{\text{female}},$$

where E_{female} is the orthogonal matrix of eigenvectors of $\boldsymbol{K}_{\text{female}}$ and where Λ_{female} is the diagonal matrix of eigenvalues of $\boldsymbol{K}_{\text{female}}$. Likewise, we can simulate the variance and covariance of the interattribute distances of our aligned male face set with μ_{male}:

$$\begin{bmatrix}
-37.90 \\
-38.46 \\
-17.83 \\
-18.11 \\
31.80 \\
93.46 \\
-56.25 \\
56.31 \\
-49.86 \\
49.86 \\
0.57 \\
-0.62
\end{bmatrix}$$

and $\boldsymbol{K}_{\text{male}}$:

$$\begin{bmatrix}
9.38 & 5.89 & -3.98 & -5.03 & -3.31 & -2.74 & -3.35 & 4.47 & -2.56 & 4.29 & -1.99 & -1.06 \\
5.89 & 10.63 & -5.11 & -4.98 & -3.43 & -2.67 & -5.82 & 4.61 & -3.14 & 1.40 & 1.16 & 1.69 \\
-3.98 & -5.11 & 7.29 & 5.53 & -2.39 & -1.37 & 0.46 & 0.54 & -0.12 & 0.97 & -1.03 & -0.62 \\
-5.03 & -4.98 & 5.53 & 7.50 & -2.21 & -0.88 & 0.20 & -0.47 & -0.68 & -0.46 & 1.14 & 0.41 \\
-3.31 & -3.43 & -2.39 & -2.21 & 11.40 & -0.04 & 3.17 & -3.17 & 2.98 & -3.08 & 0.17 & -0.30 \\
-2.74 & -2.67 & -1.37 & -0.88 & -0.04 & 7.74 & 5.22 & -5.90 & 3.45 & -3.04 & 0.56 & -0.07 \\
-3.35 & -5.82 & 0.46 & 0.20 & 3.17 & 5.22 & 8.79 & -6.95 & 0.59 & -0.72 & -0.57 & -0.82 \\
4.47 & 4.61 & 0.54 & -0.47 & -3.17 & -5.90 & -6.95 & 8.63 & -2.06 & 0.88 & -1.05 & 0.44 \\
-2.56 & -3.14 & -0.12 & -0.68 & 2.98 & 3.45 & 0.59 & -2.06 & 6.50 & -3.47 & -1.15 & -0.30 \\
4.29 & 1.40 & 0.97 & -0.46 & -3.08 & -3.04 & -0.72 & 0.88 & -3.47 & 6.92 & -2.48 & -0.93 \\
-1.99 & 1.16 & -1.03 & 1.14 & 0.17 & 0.56 & -0.57 & -1.05 & -1.15 & -2.48 & 5.32 & -0.18 \\
-1.06 & 1.69 & -0.62 & 0.41 & -0.30 & -0.07 & -0.82 & 0.44 & -0.30 & -0.93 & -0.18 & 1.97
\end{bmatrix}$$

A Matlab function (i.e., create_ feature_pts) implementing this method is freely available at http://mapageweb.umontreal.ca/gosselif/alignTools/)