The Invariance of Recognition to the Stretching of Faces is Not Explained by Familiarity or Warping to an Average Face

Catrina Hacker[1] and Irving Biederman[1,2]

1. Neuroscience Program, University of Southern California, Los Angeles, CA
2. Department of Psychology, University of Southern California, Los Angeles, CA
Corresponding Author: chacker@usc.edu

Abstract. Stretching (or compressing) a face by a factor of two has no effect on its recognition as assessed by the speed and accuracy of judging whether the face is that of a celebrity (Hole, 2002). This invariance has stood as a challenge to all contemporary accounts of the relation between neurocomputational measures of face similarity and face recognition. We extend the documentation of strong invariance over compression to a factor of four and show that the deformation so produced is sufficiently great that the resultant image is as similar to markedly different faces—even those differing in race, sex, and expression-- as it is to the original face. The invariance to face compression is readily witnessed with less familiar celebrities and unfamiliar faces ruling out a role of exposure to transformed images of particular faces through depth rotation or viewing pictures at varied viewing angles. We additionally discount the possibility that faces are "un-stretched" by warping them to an average face. Instead, we suggest that the percept of an elongated face provides a signal for the shrink-wrapping of receptive fields to conform to an attended object, a phenomenon witnessed in single unit activity in the macaque by Moran and Desimone (1985) which may serve, more generally, as the underlying neural mechanism for object-based attention.

Keywords: Face recognition, Invariance to compression and stretching, Gabor-jet measure of face similarity, Object-based attention, Face warping, Stretched Faces

Introduction

In 2002, Hole et. al. reported that stretching a face along its vertical (Fig. 1), horizontal or diagonal dimensions by a factor of two (1:2) had no effect on the speed or accuracy of face identification in which headshots were to be judged as famous or not famous. Hole's result has stood as a challenge to almost all computationally-based models of face representation in that the gigantic image deformation imposed by such stretching should exact a sizable cost in the capacity to identify the face. The computational cost, and the lack thereof, will be addressed in the Results section.

Figure 1. Original headshots (1:1) of Leonardo DiCaprio (left) and Morgan Freeman (right) and headshots stretched vertically (1:2) to the same degree as the stimuli in Hole's 2002 publication.

This invariance is also remarkable given the scarcity of images of stretched faces in our daily visual world. An instance in which we do encounter a face whose image has undergone stretching occurs when we view a picture of a face on a magazine cover lying flat on a low table. When viewing the face at a distance, if the chin to forehead dimension is collinear with the observer's viewpoint, the image of that dimension will be lengthened, yielding an image of the face that is stretched vertically.

A similar, although not identical, transformation occurs with rotation in depth of a face. When a face is rotated from its frontal view around its vertical axis, the aspect ratio of the face elongates vertically as the horizontal dimension foreshortens. Unlike the affine transformations used by Hole, when rotating a face some information is lost due to self-occlusion while other parts are revealed through accretion. Further, rotation in depth of a face has been shown to markedly increase the difficulty of face recognition or face matching (e.g. Biederman & Kalocsai, 1997; Liu et. al., 2005; Wallraven et. al., 2002; Favelle et. al., 2007; O'Toole et. al., 1998; Zhu, et al., 2018), while stretching has not.

Increased familiarity with a face has been shown to facilitate a number of face recognition and discrimination tasks and familiar faces may be processed differently than unfamiliar faces (Ramon et al., 2015; Schwartz & Yovel, 2016; Klatzky & Forrest, 1984; Ellis et. al., 1979; Johnston & Edmonds, 2009). Bruce (1982) demonstrated that familiarity with a face could facilitate greater invariance to rotation in depth, so it is possible that repeated exposure to familiar faces from different viewpoints and under different conditions builds a robust representation of that face which facilitates invariance to stretch. If this is the case, then invariance to stretch would be expected to be stronger for familiar than unfamiliar faces.

Here we extend the limits of this invariance to stretch by studying recognition of faces out to four times vertical stretch. Additionally, we address the possibility that face familiarity plays a role in mediating this invariance by collecting individual familiarity ratings from each subject for the famous faces used in the experiment. Finally, we offer an account for how this remarkably robust invariance to stretch could be so readily demonstrated.

Methods

Subjects

113 undergraduates from the University of Southern California, 86 female, 6 left-handed, 3 ambidextrous, mean age: 20.0, S.D. 2.2, were recruited through USC's Psychology Department subject pool and received class credit as compensation. An additional eight subjects were eliminated from analysis: three had markedly lower average familiarity ratings than the rest of the sample (selecting the lowest rating of familiarity for > 33/66 faces), two had at least one excessively high reaction time (> 9 seconds), two showed regularly patterned key responses indicating that they were unlikely to have been processing the display, and one had an overall accuracy far below chance (< 15%) indicating that they likely switched the keys when making their responses.

Stimuli

Headshots of 66 celebrities (33 female) were chosen based on high familiarity ratings in previous face recognition tasks (Hacker et al., 2018; Meschke et al., 2017). Three images of each celebrity were obtained via Google image search. The images varied in pose, hairstyle, expression, lighting , etc., and each was randomly assigned to one of the following three conditions: a) the original (unchanged) image, b) compressed to either two or c) four times its original width. Additionally, images of 198 different non-celebrities (99 female) were obtained through a Google image search and randomly assigned to one of the same three levels of horizontal compression as the celebrities: 1:1, 1:2 or 1:4 (Fig. 2). Apparent stretching was achieved by compressing the horizontal dimension to change the aspect ratio of the headshot while holding constant the maximum vertical dimension of the images rather than increasing the vertical dimension and holding the horizontal dimension constant. This differs from Hole's original stimulus manipulation but is consistent with other studies exploring extreme variations in degrees of stretching (e.g., Gilad-Gutnick et. al., 2018). Compressing (vs. stretching) thus restricted the longest dimension to its original extent, reducing the likelihood of face features being projected to less sensitive parafoveal regions but at a possible cost of a reduction in resolution along the compressed dimension. The quality of celebrity images did not differ between celebrities and non-celebrities as evidenced by near-chance performance on trials containing a celebrity individually rated as unfamiliar. Additionally, there was a strong relationship between performance and rated level of familiarity, with the familiarity ratings varying from subject to subject for a given image (Fig. 6). Each image was cropped to a square with only the face and tops of the shoulders visible before compressing. The background was then removed and replaced by a homogeneous grey background in Adobe Photoshop to remove any background cues that might be diagnostic of celebrity status. All images were then grey-scaled to eliminate any predictive information in the colors of the images.

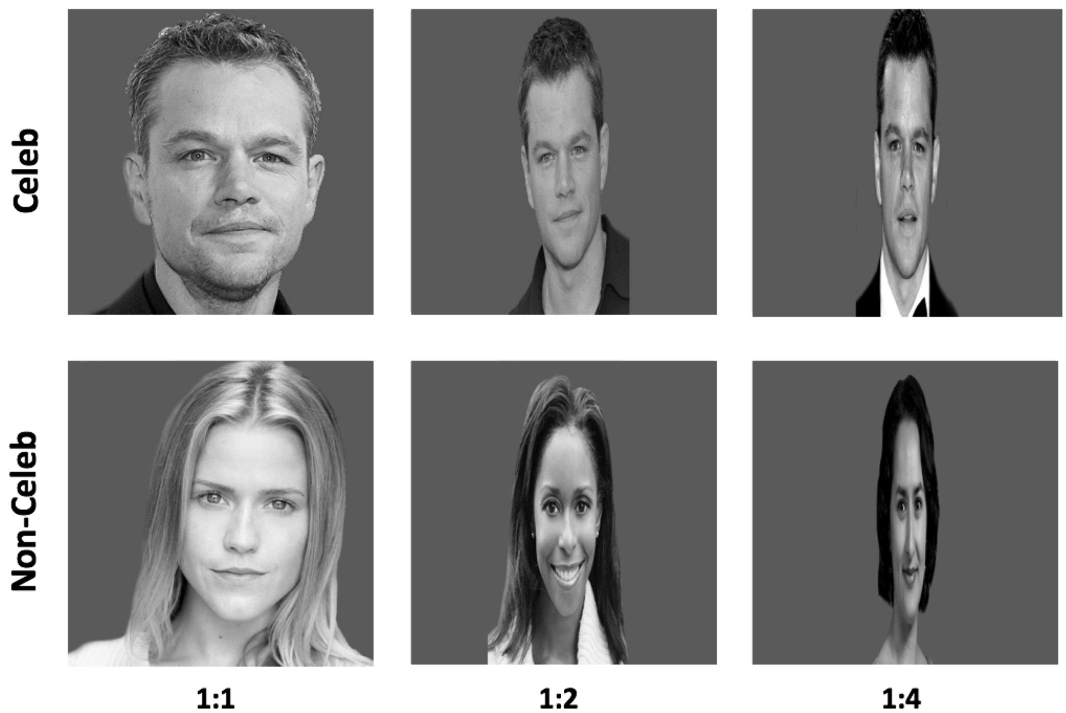

Figure 2. Top row: One sample celebrity identity (Matt Damon) under each level of compression. Note that while the same celebrity appeared at all three levels, different images were used to prevent learning of the picture rather than recognition of an identity. Bottom row: Images of three non-celebrities; one image under each degree of compression. Non-celebrity identities were not repeated across conditions. Images shown here were not used in the experimental task for copyright reasons, however they are representative of experimental stimuli.

Experimental Task

The task can be taken on testable.org (testable.org/t/373693470). Subjects viewed one headshot at a time and indicated via key press, "as quickly and as accurately as possible," whether the image was that of a celebrity or not. Each of the three headshots of a given celebrity was viewed once under each level of compression, while each non-celebrity's headshot was viewed only once with an equal number under each condition of compression for a given subject (Fig. 2). Subjects viewed a given celebrity headshot only once under a single level of compression to prevent learning of the *picture* rather than recognition of a face. Headshots were randomly assigned to a level of compression, and the same headshots appeared under the same level of compression for all subjects. After completing five practice trials, subjects performed four blocks of 99 trials, half of which contained an image of a celebrity. The experiment was balanced within subjects and stimuli from each of the six conditions occurred randomly in permutations of six.

After completing the task, subjects rated their familiarity with the faces of the 66 celebrities, designated by name, on a scale of 1 to 5 with 1 = "unfamiliar" and 5 = "highly familiar". A mean of 54.9% of the celebrities were rated as a "5" (Fig. 3) and were assigned to a high familiarity bin; ratings of a 2, 3 or 4 were sorted into a single low familiarity bin. Celebrities rated as a "1" were excluded from all subsequent analyses for that subject although they were included in the group data shown in Fig. 6. It is possible that the faces of some celebrities whose name was rated as a "1" were recognized as familiar when judging the faces in that the subject might have been unfamiliar with the celebrity's name but not his or her face. All non-celebrity trials were processed as a third, separate category. If a given subject had fewer than three celebrities in a given familiarity bin their score was not included in the mean calculation for that bin.

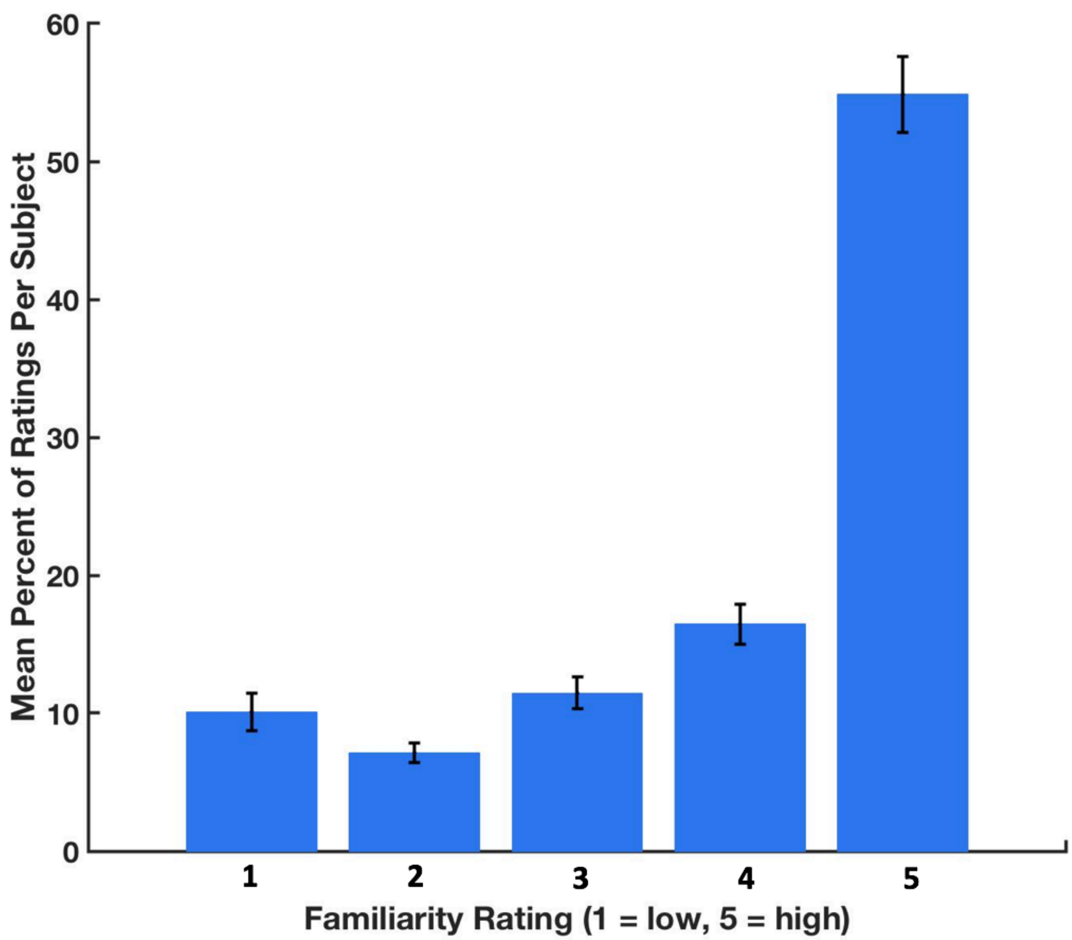

Figure 3. Distribution of subjects' familiarity ratings for the 66 celebrities. Error bars are the standard error of the mean.

Gabor Jet dissimilarity scaling of stretched faces

The Gabor Jet model is a model based on V1 simple-cell, hypercolumn filtering which yields a similarity value for pairs of faces that almost perfectly predicts their psychophysical discriminability. Correlations between Gabor dissimilarity values and error rates on a match to sample task are in the mid .90s (Yue et. al., 2012). Lest it be thought that the actual image changes produced by compression are small with respect to the representation of faces, we employed the Gabor Jet model to scale the magnitude of the difference between compressed and un-compressed versions of the same image relative to the difference between two people with faces of highly different appearance. Each of the three images (at different levels of compression) of a celebrity's face was centered within a grid of 100 Gabor jets, each containing five scales, eight orientations and two phases (Fig. 4).

1:1 **1:2** **1:4**

Figure 4. Placement of the 10x10 Gabor jet grid on one sample image of Will Smith under all three conditions of stretch (un-stretched or stretched by a factor of two or four). Each dot designates the position of one jet and each of the 100 jets represents the Gabor filtering of one simplified V1 hypercolumn with five scales, eight orientations and two phases. Each face is thus represented by a vector of 8,000 values.

The higher the Gabor dissimilarity value for a pair of faces, the more distinguishable the faces (Yue, et al., 2012). The magnitude of the change imposed by compressing a face by a factor of two or four was substantial: it proved to be as high as comparing that same face to a very different appearing person. For example, the dissimilarity of Will Smith's face to his face compressed by a factor of 4 is slightly greater than the dissimilarity between the uncompressed faces of Will Smith and Angelina Jolie who also differed in both clothing and expression (Fig. 5).

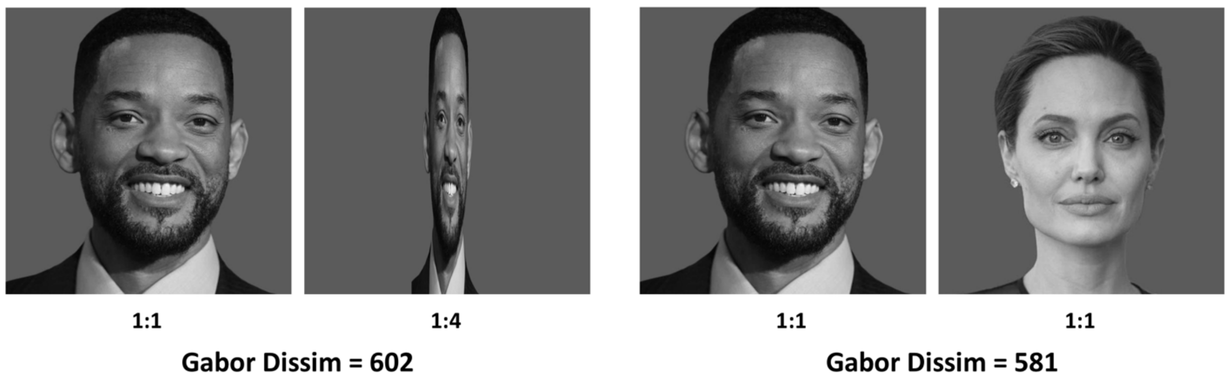

Figure 5. An illustration of the comparable dissimilarity between a single image of a face (Will Smith in this example) un-compressed and compressed by a factor of four (left), or compared to the face of a different person (Angelina Jolie, right). The magnitude of the dissimilarity between the original picture of Will Smith and that same image compressed by a factor of four is slightly larger (i.e., more dissimilar) than the dissimilarity between Will Smith and Angelina Jolie's uncompressed faces which are readily distinguished.

Results

Accuracy of judging celebrity status as a function of the rated familiarity of that celebrity

Accuracy increased almost linearly by about 10% with each point increase in the individual rated level of familiarity from slightly above chance (for "1s") to 93.8% (for "5s") (Fig. 6). Accuracy on trials rated as a "1" was at 57.5%, a value near, although reliably above, chance and with a small effect size, $t(112) = 3.13$, $p < .01$, $d = 0.29$. As noted earlier, it is possible that some of these cases were ones where the subject had encountered the face but did not know the name. That the accuracy of judging unfamiliar faces was near chance and the strong relationship between accuracy and familiarity provide evidence that there were no underlying cues in the celebrity images which would have allowed a differentiation of celebrity from non-celebrity headshots without recognition of the faces.

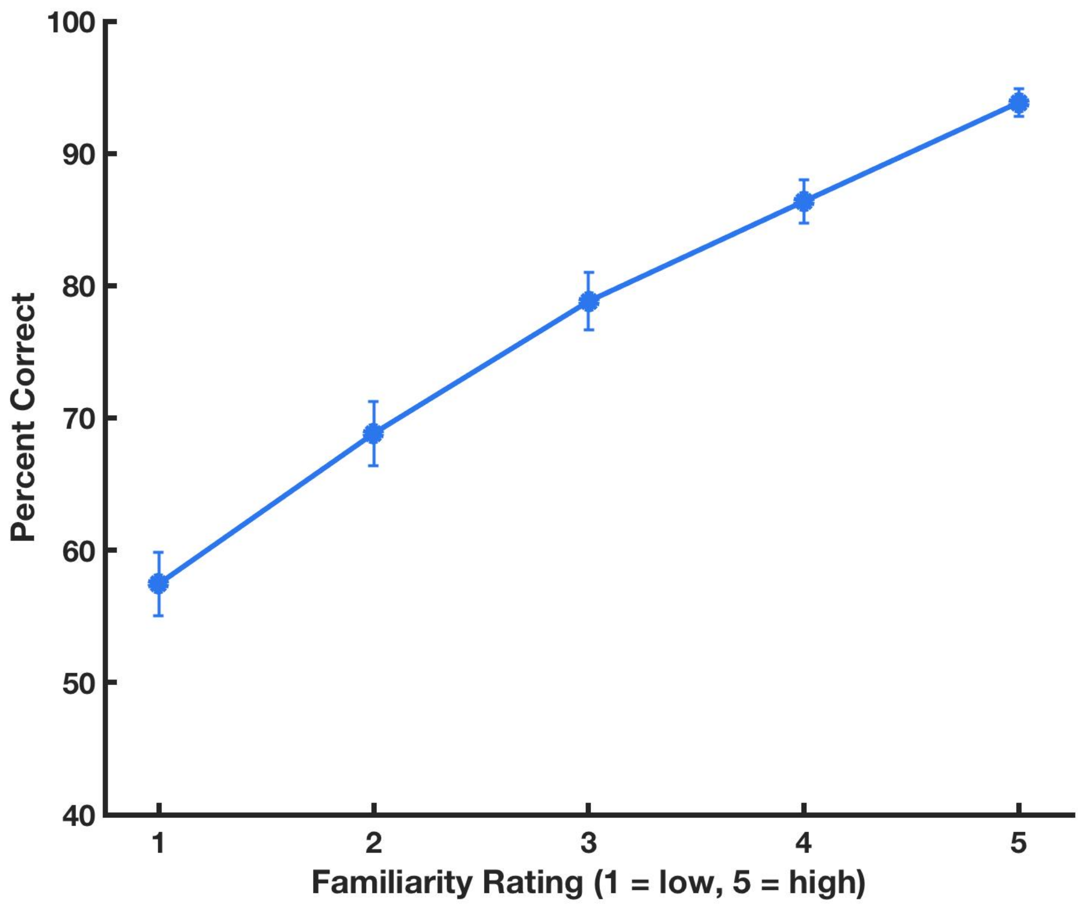

Figure 6. Percent correct judgment of celebrity status on celebrity trials as a function of individual familiarity ratings. Error bars are standard errors of the mean.

Effects of face compression on the speed and accuracy of judgments of fame

The cost of compression on accuracy (Fig. 7) and reaction times (Fig. 8) was generally miniscule in magnitude. Compressing an image of a face by a factor of four produced an overall increase in error rates of 1.97% and an increase in RTs of 59 msec. Given the enormous power afforded by 113 subjects the slight increase in error rates and RTs were both significant but with low to moderate effect sizes as indicated by the $\eta_p{}^2$ values, Accuracy: $F(2, 194) = 7.94$, $p < 0.001$, $\eta_p{}^2 = 0.08$, RTs: $F(2, 190) = 14.67$, $p < 0.001$, $\eta_p{}^2 = 0.13$. A post-hoc Tukey test revealed that accuracy on non-celebrity trials was significantly different between original (1:1) and both two times compression, $p < 0.05$, and four times compression, $p < 0.01$. For high familiarity trials, the post-hoc Tukey analysis revealed a significant difference between reaction times on two times and four times compressed conditions, $p < 0.01$, and un-compressed and four times compressed conditions, $p < 0.001$. The post-hoc Tukey analysis also revealed significant differences between reaction times between the un-compressed and four times compressed conditions, $p < 0.001$, and the two times and four times compressed conditions, $p < 0.001$.

As will be discussed subsequently, these costs at four times compression are possibly the effects of reduced resolution of the facial features from their compression rather than difficulty in matching a compressed face to a representation of a familiar person.

Accuracy and reaction times as a function of high, low, or no familiarity

Celebrities rated as a 2, 3 or 4 were combined into a low familiarity group, while those rated as a 5 formed a high familiarity group and non-celebrities made up a third and separate group. Subjects made far fewer errors, $F(2, 194) = 48.38$, $p < 0.001$, ${\eta_p}^2 = 0.33$, and responded more quickly, $F(2, 190) = 67.23$, $p < .001$, ${\eta_p}^2 = 0.42$, in recognizing celebrities with whom they were highly familiar compared to those of low familiarity. This advantage in judging highly familiar celebrities was independent of the degree of stretch for both accuracy, $F(4, 388) < 1.00$, *ns*, and reaction times, $F(4, 380) = 2.2$, *ns*, indicating that increased exposure to a highly familiar face (affording greater variability in viewpoints) had no reliable effect on the invariance to stretch.

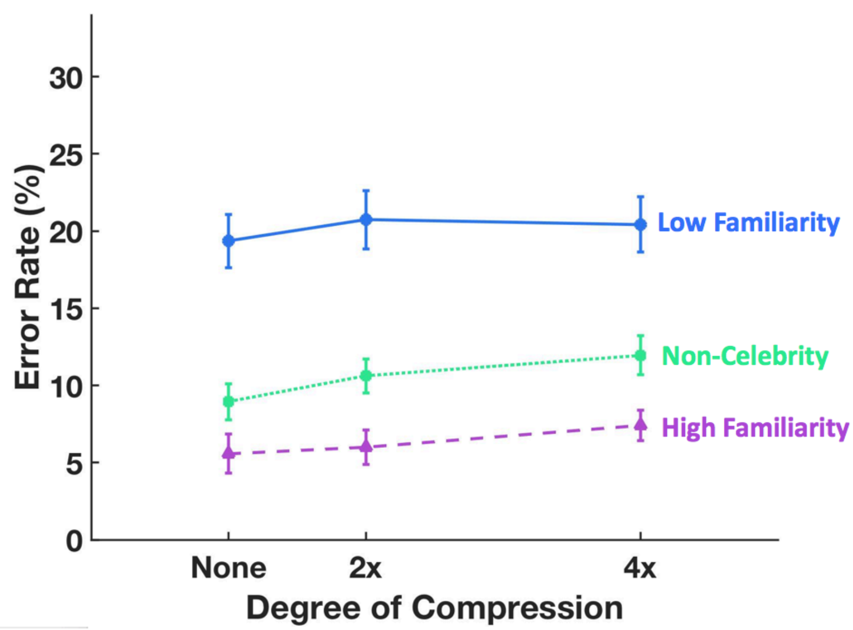

Figure 7. Error rates as a function of level of familiarity and degree of horizontal compression. Error bars are S.E. of the mean.

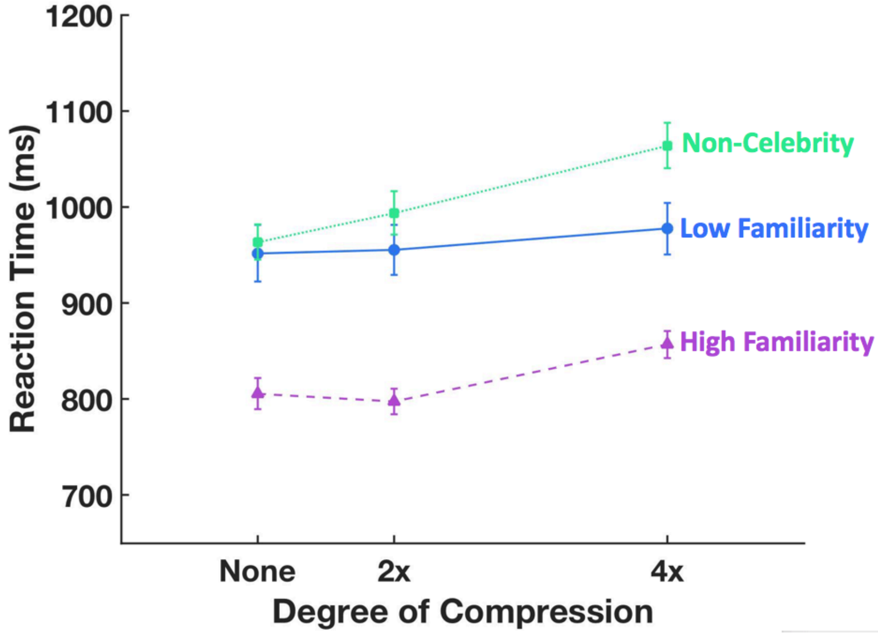

Figure 8. Reaction times as a function of level of familiarity and degree of horizontal compression. Error bars are S.E. of the mean.

Repetition of celebrity identity

The same celebrity identity was repeated three times over the course of the experiment, with a different headshot each time. Error rates declined less than 1% over the three appearances. Although RTs declined by 73.3 ms with repetition of the celebrity (and increased familiarity with the task in general), there were no interactions of repetition with the major experimental variables, and the decline could be attributed mostly to non-celebrity trials rather than celebrity trials.

Discussion

Compression of a face results, at best, in only a minuscule increase in RTs or error rates and those effects, at 4X, might be attributable to reduced resolution of the facial features because of the compression as discussed below. Gabor jet scaling of the dissimilarity between an un-compressed face and its compressed counterpart shows that this difference is enormous; as great as, if not greater than, the difference between two very different faces as noted in the example of Will Smith and Angelina Jolie (Fig. 5). Despite these large effects of compression on the similarity scaling of face images, we demonstrated strong invariance of face recognition up to a compression of four times an image's original horizontal dimension. That is, not only is recognition possible under compression, but the accuracy and speed of recognition are virtually unaffected.

Limits of invariance to stretch as a result of decreasing image resolution

Recent work from Gilad-Gutnick et. al. (2018) has demonstrated that beyond 80% compression there is a decline in the speed and accuracy of recognition. Their designation of 80% compression derives from a compression of a factor of 5 which would leave an image at 20% of its original width. Our maximum condition, four times compression, would be 75% compression by Gilad-Gutnick et. al.'s criterion. As noted earlier, the limits of the invariance to compression observed in the present study and Gilad-Gutnick et. al. may be less one of limits to the invariance of face recognition to compression per se than on limits to the perceptual resolution of the internal face features produced by the compression of the face without changing the overall area of the image. Consider Fig. 9, which shows headshots of six celebrities under varying degrees of compression. It is subjectively apparent that at higher levels of compression increased scrutiny is required to achieve recognition. But this is less a failure of the capacity to recognize a compressed familiar face than increased difficulty in perceiving the characteristics of the face. That is, the prolonged scrutiny required in attempting to recognize Obama's face may be less a consequence of compression distortion than the lower resolution of its features independent of its aspect ratio.

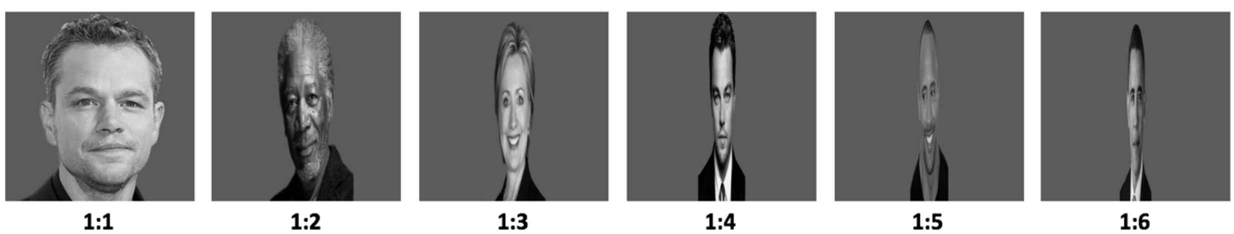

Figure 9. Images of various celebrities at increased degrees of compression in the horizontal dimension (increasing compression left to right). At the higher compression ratios, e.g., those greater than 1:4, detailed scrutiny is required to resolve a given identity. (Pictured left to right: Matt Damon, Morgan Freeman, Hillary Clinton, Leonardo DiCaprio, Dwayne Johnson, Barack Obama)

What mechanism might underlie invariance to compression?

Exposure to a rich variety of images for a given face. One possible explanation for invariance to compression is that increased exposure to familiar faces under different viewpoints and conditions generates a robust representation of that face which renders it invariant to compression. Our results do not support this hypothesis as invariance to compression was witnessed for faces that were rarely, if ever, encountered previously. The speed and accuracy of judgment of non-celebrity faces, which were completely unfamiliar, were as invariant to compression as highly familiar faces.

Warping to an average face template. Another possible explanation is that faces are "un-stretched" by matching them to an average face template, as has been described, for example, as a "faciotopy" by Henriksson et al. (2015). But this explanation is also unlikely. Warping a face to fit an average template distorts the individual shape features and their distances so much that recognition suffers. Figure 10 provides an example of the detrimental effects of warping to an average face shape, demonstrated by an algorithm developed by Kramer et. al. (2017). When the image of Obama is warped to an average face shape, the image becomes less easily recognizable—if not unrecognizable. The Gabor dissimilarity between the two images of Obama is 578, almost equal to the value of 581 between Will Smith and Angelina Jolie (Fig. 5), with the latter pair also differing in expression and clothing. Consistent with the importance of preserving shape information in recognizing faces, Russell et. al. (2007) showed that maintaining all pigment information within a face while averaging shape information led to significantly lower recognition accuracy. The configural effect, an important aspect of face recognition, also supports the importance of the original shape of a face, as it demonstrates that small changes to the shapes or relative positions of individual face parts affects the identifiability of a face (Tanaka & Farah, 1993; Xu et. al., 2014). A face template built up from an average of many faces may play more of a role in face *detection* than in face recognition.

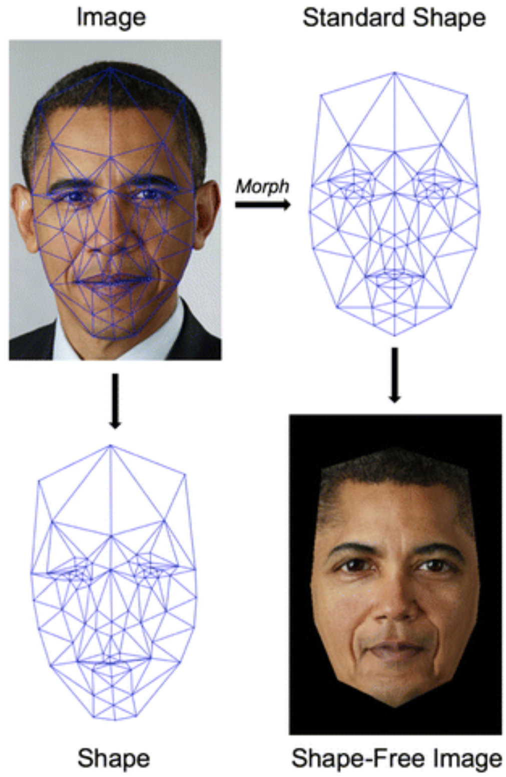

Figure 10. Figure 1 from Kramer et. al. (2017) showing the effects of warping a face to an average face shape. The shape-free image of Obama (bottom right) is not as easily identifiable as the original image of Obama (top left). The Gabor dissimilarity between the two images of Obama is 578, which is almost as large as the dissimilarity of 581 between Will Smith and Angela Jolie (Fig. 5).

Attentional modulation of receptive fields (r.f.s) as an explanation of the invariance of face recognition to the compression of faces. There is evidence that FFA retains aspects of the spatial, Gabor-like tuning, characteristic of early visual cortex (Yue et al., 2006), albeit with larger r.f.s than in earlier visual areas (e.g., Witthoft et al., 2016). LOC, by contrast, appears to maintain an edge-based type of representation, closely equivalent to a line-drawing, e.g., Grill-Spector et al. (2001). Large, overlapping receptive fields in face selective areas can generate the configural effects which serve to preserve and amplify the impact of the fine metric differences that allow non-prosopagnosics to distinguish highly similar faces (Xu et al., 2014).

We suggest that the attentional modulation of receptive fields, as first demonstrated by Moran and Desimone's (1985) single unit recordings in V4 of the macaque, and confirmed under more general conditions by Reynolds et al. (2002), may be sufficient to yield invariance to compression. Specifically, these investigators showed that a normally effective stimulus for a V4 cell, for example a vertical red bar within the receptive field of the cell, would not elicit heightened activity if the monkey was not attending to that region of the cell's receptive field. It was as if the attentional cueing caused the cell to "shrink wrap" so that the cell was responsive only to the attended area. The percept of a compressed face may provide a signal to undertake a similar compression of receptive fields, perhaps in face-selective cortex. Such compression of receptive fields may be sufficient to eliminate differences between uncompressed and horizontally compressed faces.

Whatever its role in the achievement of invariance of face recognition to image transformations, the deformation of r.f.s may be critical to normal face recognition. Given the evidence for spatial tuning in FFA, there has to be some restriction so that the r.f.s are only tuned to regions *inside* the face, otherwise the context would impose, in general, massive irrelevant inputs to the cells coding a face, particularly those with large r.f.s along the perimeter of the face.

More generally, the modulation of receptive field shape as documented by Moran and Desimone (1985) may be the underlying neural mechanism for object-based attention (Scholl, 2001; Mueller & Kleinschmidt, 2003; Behrmann et. al., 1998; Marino & Scholl, 2005). Recent work by Kay et. al. (2015) has demonstrated that the size of population receptive fields (pRFs) can be attentionally modulated in response to task demands. This result joins a growing body of work demonstrating that receptive fields in extrastriate cortex are fluid and flexible, offering a possible explanation for the invariance of recognition to compressed faces.

Conclusion

Consistent with Hole's (2002) original discovery of strong invariance in the recognition of compressed faces, we showed that compression up to four times the original dimensions incurred at best little cost in the speed or accuracy of recognition. There was no interaction between level of familiarity of a face and its invariance to compression indicating that the invariance was not dependent on prior exposure to a particular face. Nor could the invariance be explained by the warping of a compressed face to an average face, as such warping produces faces that are less easily recognizable than the original. We suggest that the strong invariance to compression when recognizing faces is a special case of object-based attention whereby the percept of an elongated face elicits a corresponding deformation of receptive fields in face-selective areas allowing for the recognition of a face to proceed independent of the compression.

Acknowledgements

Supported by NSF BCS 0617699 and the Dornsife Research Fund. We thank Emily X. Meschke for helpful discussions and Sarah B. Herald for both her helpful discussions and her calling our attention to the problems associated with morphing to an average face.

References

Behrmann, M., Zemel, R.S. & Mozer, M.C. (1998). Object-based attention and occlusion: evidence from normal participants and a computational model. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1011-1036. doi: 10.1037/0096-1523.24.4.1011.

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society London: Biological Sciences*, 352, 1203-1219. doi: 10.1098/rstb.1997.0103.

Bruce, V. (1982). Changing faces: visual and non-visual coding processes in face recognition. *British Journal of Psychology*, *73*, 105-116. doi: 10.1111/j.2044-8295.1982.tb01795.x.

Ellis, H.D., Sheperd, J.W. & Davies, G.M. (1979). Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. *Perception*, *8*, 431-439. doi: 10.1068/p080431.

Favelle, S. K., Palmisano, S. & Maloney, R. (2007). Things are looking up: Differential decline in face recognition following pitch and yaw rotation. *Perception*, *36*, 1334-1352. doi: 10.1068/p5637.

Gilad-Gutnick, S., Harmatz, E.S., Tsourides, K., Yovel, G. & Sinha, P. (2018). Recognizing Facial Slivers. *Journal of Cognitive Neuroscience*, *30,* 951-962. doi: 10.1162/jocn_a_01265.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41, 1409-1422.

Hacker, C.M., Meschke, E.X. & Biederman, I. (2018). A Face in a (Temporal) Crowd. *Vision Research*. doi: 10.1016/j.visres.2018.02.007.

Hole, G., George, P., Eaves, K. & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception*, *31*, 1221-1240. doi: 10.1068/p3252.

Johnston, R.A. & Edmonds, A.J. (2009). Familiar and unfamiliar face recognition: a review. *Memory*, *17*, 577-596. doi: 10.1080/09658210902976969.

Klatsky, R.L. & Forrest, F.H. (1984). Recognizing familiar and unfamiliar faces. *Memory & Cognition*, *12*, 60-70. doi: 10.3758/BF03196998.

Kay, K. N., Weiner, K. S. & Grill-Spector, K. (2015). Attention reduces spatial uncertainty in human ventral temporal cortex. *Current Biology*, 25, 595-600. doi: 10.1016/j.cub.2014.12.050.

Kramer, R., Jenkins, R. & Burton, M. (2017). InterFace: A software package for image warping, averaging, and principle components analysis. *Behavior Research Methods*, 49, 2002-2011. doi: 0.3758/s13428-016-0837-7.

Henriksson, L., Mur, Marieke, & Kriegeskorte, N. (2015). Faciotopy—a face-feature map with face-like topology in the human occipital face area. Cortex, *72*, 156-167. doi: 10.1016/j.cortex.2015.06.030.

Liu, C.H., Collin, C.A., Farivar, R. & Chaudhuri, A. (2005). Recognizing faces defined by texture gradients. *Perception & Psychophysics*, *67*, 158-167. doi: 10.3758/BF03195019.

Marino, A.C. & Scholl, B.J. (2005). The role of closure in defining the "objects" of object-based attention. *Perception & Psychophysics, 67*, 1140-1149. doi: 10.3758/BF03193547.

Meschke, E., Hacker, C.M., Juarez, J., Maarek, R., Biederman, I. (2017). Can familiar faces be negatively detected at RSVP rates? *Journal of Vision*, *17*, 1027. doi: 10.1167/17.10.1027.

Moran, J. & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*, 782-784. doi: 10.1126/science.4023713.

Mueller, N.G. & Kleinschmidt, A. (2003). Dynamic interaction of object-and space-based attention in retinotopic visual areas. *Journal of Neuroscience*, *23*, 9812-9816. doi: 10.1523/JNEUROSCI.23-30-09812.2003.

O'Toole, A.J., Edelman, S. & Bülthoff, H. H. (1998). Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research, 38*, 2351-2363. doi: 10.1016/S0042-6989(98)00042-X.

Ramon, M., Vizioli, L., Liu-Shuang, J., & Rossion, B. (2015). Neural microgenesis of personally familiar face recognition. *PNAS*, 112 (35), E4835-E4844. doi: 10.1073/pnas.1414929112.

Reynolds, J. H., Pasternak, T., & Desimone, R. (2002). Attention increases sensitivity of V4 neurons. *Neuron,* 76, 703-714.

Russell, R., Biederman, I., Nederhouser, M, & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research, 47,* 157-165. doi: 10.1016/j.visres.2006.11.002.

Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, *80*, 1-46. doi: 10.1016/S0010-0277(00)00152-9.

Schwartz, L. & Yovel, G. (2016). The roles of perceptual and conceptual information in face recognition. *Journal of Experimental Psychology: General*, *145*, 1493-1511. doi: 10.1037/xge0000220.

Tanaka, J.W. & Farah, M.J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, *46A*, 225-245. doi: 10.1080/14640749308401045.

Wallraven, C., Schwaninger, A., Schuhmacher, S., Bülthoff, H. H. (2002). View-based recognition of faces in man and machine: re-visiting inter-extra-ortho. *Lecture Notes in Computer Sciences*, *2525*, 651-660. doi: 10.1007/3-540-36181-2_65.

Witthoft, N., Poltoratski, S., Nguyen, M., Golarai, G., Liberman, A., LaRocque, K. F., Smith, M. E., Grill-Spector, K. (2016). Reduced spatial integration in the ventral visual cortex underlies face recognition deficits in developmental prosopagnosia. *bioRxiv, d*oi: http://dx.doi.org/10.1101/051102.

Xu, X., Biederman, I. & Shah, M.P. (2014). A neurocomputational account of the face configural effect. *Journal of Vision*, *14*, 1-9. doi: 10.1167/14.8.9.

Yue, X., Tjan, B., & Biederman, I. (2006). What makes faces special? *Vision Research*, 46, 3802-3811.

Yue, X., Biederman, I., Mangini, M. C., von der Malsburg, C., & Amir, O. (2012). Predicting the Psychophysical Similarity of Faces and Non-Face Complex Shapes by Image-Based Measures. *Vision Research. 55*, 41-46. doi: 10.1016/j.visres.2011.12.012. doi: 10.1016/j.visres.2011.12.012.

Zhu, T., Nelken, M., Hacker, C.M., Meschke E.X., & Biederman, I. (2018). Matching Depth-Rotated Faces at Varying Degrees of Physical Similarity. Poster presented at the Annual Meeting of the Vision Sciences Society, St. Petersburg Beach, Fl. May.