Title
Modeling the Other-race Advantage with PCA

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Authors
Cottrell, Garrison W.
Haque, AfmZakaria

Publication Date
2005

Peer reviewed

Modeling the Other-race Advantage with PCA

AfmZakaria Haque (ahaque@cs.ucsd.edu)
Department of Computer Science & Engineering, 9500 Gilman Dr.
La Jolla, CA, 92093

Garrison W. Cottrell (gary@ucsd.edu)
Department of Computer Science & Engineering, 9500 Gilman Dr.
La Jolla, CA, 92093

Abstract

The Other-race effect (ORE) refers to the well-known phenomenon of people being less accurate in recognizing faces of a different race. One popular hypothesis is that we learn to use face-features that are optimal for individuating faces of our own race; thus reducing the recognition accuracy for faces of a different race. However this hypothesis has not been able to explain some advantages other-race faces have in certain tasks. For example, some recent experiments showed that in a visual search task other-race faces are found faster than same race faces when the subjects show the ORE. A race based feature selection hypothesis where deviation from the familiar race is treated as an explicit part of the encoding has been proposed to explain this other-race advantage. In this paper, we argue that the other-race advantage can be explained without this assumption. We present the results from our simulations that suggest that the other-race advantage is an inherent characteristic of an optimal feature selection model.

Introduction

It has long been known that people recognize faces from their own racial group with greater accuracy than faces from another racial group. This is known as the other-race effect (ORE), cross-race effect or own-race bias. Several meta-analyses of large number of studies in face recognition have found a strong ORE (Bothwell, Brigham, & Malpass, 1989; Shapiro & Penod, 1986).

There is naturally a strong agreement that the ORE is somehow caused by the learning history of each individual. Chance et al (1982) found that six year old children did not show a significant other-race effect; but for older participants the degree of the ORE increased with age. Although there is no conclusive answer to the question of how experience with faces can cause the ORE, there are two dominant hypotheses.

What Causes the ORE

Feingold (1914) suggested that other things being equal, the other-race effect depends on the contact with people of the other race. However, several studies (Shepard, 1981; Valentine, Chiroro, & Dixon, 1995) argued against this contact hypothesis. They suggested that contact for individuation could be more significant than mere contact. It still leaves the question that how the need for individuation can cause the ORE.

In this section we will discuss two dominant hypotheses of the ORE. Both of them concentrate on the feature selection scheme humans use for processing faces of same and different races.

Optimal Feature Selection

Optimal feature selection treats the ORE as people's inability to generalize their feature selection from the same race faces to the other-race faces. People select face-features that are optimal for identifying each individual. Since generally people are most exposed to faces of their own race, their feature selection is biased towards optimizing the recognition of this class of faces. Assuming faces of different races vary along different dimensions, their feature selection does not capture the variations of other-race faces well, reducing their accuracy in other-race face recognition.

In some ways this hypothesis is a closely related with perceptual expertise. The hypothesis can be interpreted as an effect of specialization for same race faces within the domain of face expertise (Tanaka et al., in press).

Several studies (O'Toole et al., 1991, 1994) have used Principal component analysis (Jolliffe, 1986) to model optimal feature selection. PCA finds a linear transformation to a new set of dimensions that maximizes the variance of the data. Linearity might be an undesirable constraint, but the simplified computation is well worth the compromise. Moreover PCA is a neurobiologically plausible means of feature selection since simple networks employing hebbian learning can learn to extract equivalent features (Sanger, 1989). In our simulation we modeled optimal feature selection with PCA.

Race Dependent Feature Selection

The race dependent feature selection hypothesis assumes an asymmetric feature selection scheme for faces. Levin (2000) proposed that for other-race faces race specifying information is encoded at the expense of individuating information. Loss of individuating information reduces recognition accuracy for other-race faces, causing the ORE.

In this hypothesis, for other-race faces, feature selection is optimal for classification by race and not for identity. Race specifying information, which can be treated as the

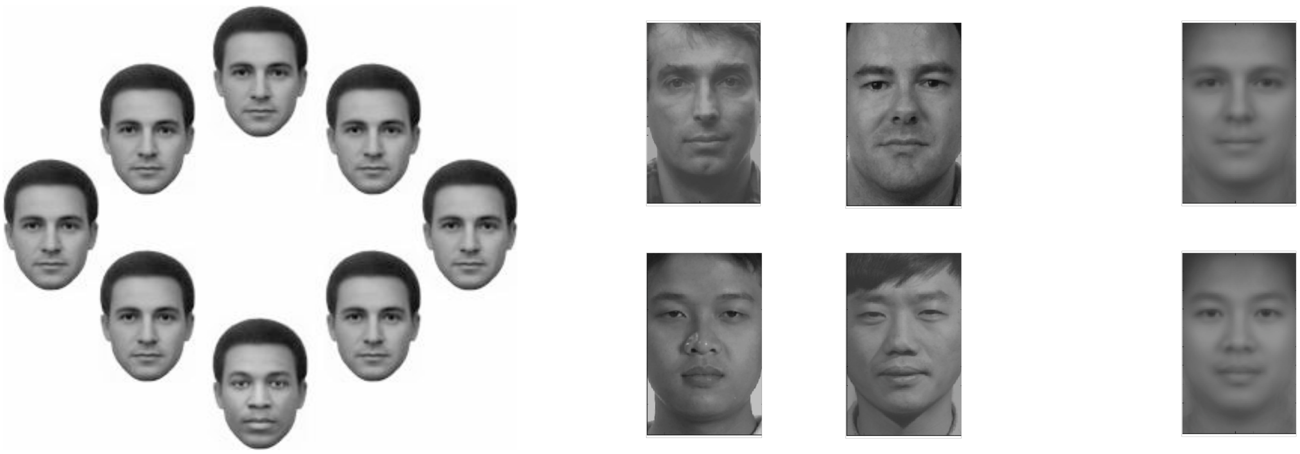

Figure 1: Stimuli. Visual search task [7] on left, our sample stimuli in the middle, our average faces on right.

deviation from the own race is treated as an explicitly encoded feature.

Although a race dependent feature selection scheme may seem to be less intuitive, it can explain the other-race advantage we discuss in the next section.

Visual Search Asymmetry Favoring Other-race Faces

Levin (2000, 1996) found that people who show the ORE are significantly faster in searching for an other-race face among same-race faces than the reverse. The stimuli (Figure 1) consisted of one White average and one Black average face. The faces were processed to have identical skin shading, hair, ears and jaw lines and differed only in internal features.

Triesman & Gormican (1988) showed that visual search for feature positive target among feature-negative distracters is faster than the reverse. This effect, called the visual search asymmetry, was assumed to occur since the feature-positive target stands out among the feature negative distracters. But feature positive distracters effectively hide the feature negative target in noise, making it harder to detect.

Levin (2000) applied Triesman & Gormican's feature positive idea by suggesting that the search asymmetry favoring other-race face can be explained if other-race faces are more feature-positive. In race dependent feature selection, other-race faces are encoded with race specifying information, which naturally makes them feature positive. Thus, this hypothesis fits well with the search asymmetry.

The possibility of other-race faces being more feature-positive in an expertise model has never been explored. If we use PCA for feature selection, an analogous idea of feature positiveness is the amount of "surprise" that the encoding of a face induces. We can think of this as a measure of the mismatch between the face and the internal representational space of the model. We will use the information content of a random variable to model this. We describe this idea further in our second simulation. Targets with higher information content will provide more clues to the visual system for search while distracters with higher information content would hide the target in noise. In next two sections we describe our simulations.

Simulation 1: the Other-race Effect

Background

A typical human experiment demonstrating the ORE is designed as a standard recognition task. Participants see a study set S of faces where half of the faces are from their own race and half from a different race. Then they are shown another set N of faces, half of which are from the previous set. For each face, the participants have to say if it is from the study set. From their response, the discriminability score d' is computed. A significantly greater d' for same race faces reflects the other-race effect (O'Toole et al., 1994).

The concept of "own race" can be modeled with a training set containing a large proportion of faces of one race. O'Toole et al. (1991) used Principal Component Analysis on a dataset with 95% Caucasian and 5% Japanese faces. They defined features as a subset of the principal components on the training set and the probability of recognizing a novel face as the cosine between a face and its reconstruction from the representation space. They found that novel other-race faces had a higher d' than novel same-race faces.

In this experiment we extended the O'Toole et al. (1991) work by adding a recognition memory component to the model. In particular, we used the Generalized Context Model (Nosofsky, 1986, 1988) in the representation space to model recognition memory. We simulated a typical human experiment with our model and found a significantly strong ORE.

Table 1: The other-race effect simulation results

Majority racial group	p_{mean}(hit)	$p_{mean}(FA_{Caucasian})$	$p_{mean}(FA_{Asian})$	$\Delta d'_{mean} = d'_{mean,Caucasian} - d'_{mean,Asian}$	significance (p-value) of $\Delta d'$
Caucasian	.69	.21	.42	.58	<< .05
Asian	.70	.22	.50	-.73	<< .05
None	.73	.30	.32	.3	~.37

Model

Similar to O'Toole et al (1991, 1994) we used PCA on a training dataset to model the learning of feature selection by long-term experience. A subset of the eigenvectors with the largest eigenvalues was used as features (Turk & Pentland, 1991). The recognition memory was modeled with a version of GCM (Nosofsky, 1986, 1988; Dailey, Cottrell, & Busey, 1999) where, given the representation of a face x and a set S of already seen faces, the probability of recognition is

$$p(x \in S) = \beta \sum_{y \in S} e^{-d_{x,y} / \sigma}$$

Here β linearly normalizes the summed similarity to a probability. $d_{x,y}$ is the Euclidian distance between the principal component representation vector of x and y . σ determines how much a learned representation contributes to recognition. σ should be of the order of d to keep the exponential term in a reasonable range. In this paper we report the results obtained by setting σ to twice the minimum d between study faces.

A forced choice yes/no recognition procedure can be modeled by responding yes if p exceeds a criterion γ. The optimal criterion would be the mean of the distribution of p for new and old faces. Signal detection methodology maps easily onto this Yes/No task since the distribution of p for old faces can be thought of as the signal distribution and the distribution of p for new faces as the noise. Old faces with p greater than γ are considered hits and new faces with p greater than γ are considered false alarms. A d' score can be computed in the standard way. A significantly lower d' for other-race faces will show the other-race effect.

Stimuli

Our stimuli consisted of 64 Caucasian and 64 Asian 128 x 192 gray scale face images extracted from FERET database (Phillips, et al., 1997) release 2. The face images were cropped and linearly warped so that the eye and mouth positions line up across images. They were also normalized for brightness and contrast. Some sample stimuli are shown in Figure 1.

Method

1. In the first experiment, our training set for PCA, T contained 44 Caucasian and 4 Asian randomly chosen face images. We kept 20 of the eigenvectors with largest eigenvalues. A set, S containing 10 Caucasian and 10 Asian face images (randomly chosen and different from T) was used as the study set. N, a set of 20 Caucasian and 20 Asian face images (different from T and superset of S) was used as test set. The same simulation was done 50 times with randomly chosen datasets.
2. We ran the same experiment by switching the majority and minority race. In this experiment, T contained 44 Asian and 4 Caucasian faces images.
3. As a control group we ran the same experiment with unbiased learning history. In this experiment, T contained 24 Caucasian and 24 Asian face images.

Results & Discussion

As Table 1 shows, in first two experiments we found a strong and significant (p << .05) bias (Δd') favoring other-race face recognition. In the third experiment, where the learning history was not biased towards any race, there was no significant difference in the discriminability score of any one race. This is essentially the classic other race effect. Figure 2 shows the ROC curve for the Caucasian majority and the control group experiment.

Simulation 2: Visual Search Asymmetry

Motivation

In this experiment we explore the feature positiveness of faces in our optimal feature selection model. In the previous experiment we found that when the learning history has a large proportion of faces from one race, our PCA-based model shows the other-race effect. In this experiment we found that under such biased learning history the other-race faces are more feature positive than the same race faces.

Model

Shannon Entropy, $-\int p_x \log(p_x) dx$ of a random variable x is often treated as the expected information content of x (Shannon, 1948; MacKay, 2003). Information content, $-\log(p_x)$, in some sense measures how much of an outlier a given value of a random variable is. Feature positiveness essentially says how much activation a stimuli causes in the feature detectors. If the distribution in representation space is zero-mean unimodal, then feature positiveness is similar to how much of an outlier a stimulus is in representation space. A probabilistic interpretation of PCA assumes a Gaussian distribution for the latent variables (Tipping & Bishop, 1999) and PCA by definition zero-means the data.

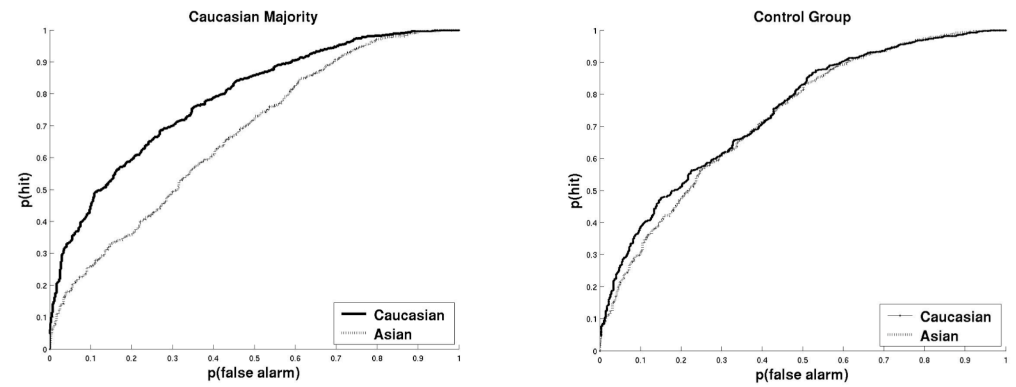

Figure 2: ROC curve. On left Caucasian majority simulation; on right control group simulation

The representation of a face in our model can thus be thought of as a multidimensional continuous random variable with a Gaussian distribution. If we use PCA for feature selection, a natural analogy for feature positiveness would be the information content of the representation of the face measured against this distribution.

For a face x and its d dimensional representation y, the information content $h(x)$ is,

$$h(x) = -\log(p(y))$$

Here $p(y)$ is the distribution of the faces people are familiar with, i.e., the faces used for PCA. Our limited data is an obstacle in estimating $p(y)$. However, the distribution can be simplified under the Gaussian assumption. Since PCA decorrelates the latent variables and uncorrelated Gaussians are statistically independent, we can decompose $p(y)$ as,

$$p(y) = \prod_{d} p(y_d)$$

Therefore,

$$h(x) = -\sum_{d} \log(p_d(y_d))$$

Since the Gaussian assumption was crucial in estimating $p(y)$ we used Kolmogorov-Smirnov test (Chakravarti, Laha, & Roy, 1967) for goodness of fit to test if the face images used for PCA in fact have a Gaussian distribution in representation space. We dropped from consideration any simulation where the test failed at 5% significance level (which occurred in 6% experiments). For the valid experiments, we approximated p(y) with a Gaussian distribution and computed the information content of novel faces. Our prediction was that in this model, minority faces would have significantly more information content than majority faces.

Stimuli

We used the same stimuli as the previous experiment. To generate prototypical stimuli for each race, we averaged same number of Caucasian and Asian faces not used for PCA. Two sample average faces are shown in Figure 1.

Method

1. We used the same stimuli as the previous experiment. Similar to the previous experiment, we use PCA on 44 Caucasian and 4 Asian face images and kept 20 eigenvectors with largest eigenvalues to develop the representation space. 40 new Caucasian and Asian face images (20 each) were projected to the representation space. The total information content of each face, h was computed. The experiment was run 50 times with randomly chosen face images.
2. We ran the same experiment by switching the majority and minority race. In this case, 44 Asian and 4 Caucasian face images were used for PCA.
3. As a control group, we used 24 Caucasian and 24 Asian face images for PCA and ran the same experiment 50 times.
4. The human experiment showing the visual search asymmetry (Levin, 2000) used average faces. Therefore we repeated the above three experiments with the test set containing one Caucasian average and one Asian average face.

Results

In the first experiment, the average information content of faces of each race from the 50 runs were used in a t test with Caucasian faces having the same information as the alternate hypothesis. The average information for Asian faces was significantly ($p \sim 10\text{-}6$) higher than that of Caucasian faces. We found similar results for Asian majority experiment where Caucasian faces contained significantly more information ($p \ll .05$). For the control group, the t test accepted the null hypothesis that both races have the same amount of information with $p \sim .26$.

We found similar results with the average faces. The average minority face contained significantly more information when the learning history was biased towards one race. For the control group, the difference between information content of the average faces was not statistically significant (p ~ .19)

Discussion

In this experiment we showed that in the expertise based hypothesis, in the presence of the ORE, the other-race faces are encoded with more information than same-race faces. This is equivalent to other-race faces being more feature positive. Using the argument developed in previous studies (Levin, 2000), as mentioned above, this explains the visual search asymmetry.

Although we tested our Gaussian assumption with Kolmogorov-Smirnov test, assuming a functional form for the density may be too restrictive. We found qualitatively similar results using kernel density estimation. However, the small number of number of data points compared to the dimensionality of the representation space makes the nonparametric estimation less reliable.

A natural question is why other-race faces contain more information in our model. Since the dataset used for PCA had a large number of majority-race faces and PCA zero-means the data, the new minority-race faces ended up further away from the majority faces in representation space. This effectively lowered their probability and increased information content.

Other-race faces containing more information and being less discriminable may seem paradoxical. Although other-ace faces contain more information in the representation space, as we showed in first simulation, the representation space is not optimal for recognizing them, making them less discriminable.

Future work

There are other instances of the other-race advantage (Levin, 2000) that we are not addressing in this paper. However, preliminary work shows that our model should explain those effects as well. We will be applying our model to those effects in near future.

Conclusion

In this work, we showed that our simple expertise-based model could explain some seemingly paradoxical human data. While, we are far from providing a conclusive framework for visual features people use for face-recognition, we hope this will help our understanding of this interesting domain.

Acknowledgement

We thank Daniel Levin for providing us with his stimuli. We also thank Carrie Joyce, Eric Wiewiora and other members of Gary's Unbelievable Research Unit (GURU) for earlier comments on this work. This research was supported by NIH grant MH57075 to GWC.

References

Bothwell, R. K., Brigham, J. C., & Malpass, R. S. (1989). Cross-racial Identification. Personality and Social Psychology Bulletin, 15.

Chakravarti, Laha, and Roy, (1967). Handbook of Methods of Applied Statistics, Volume I, John Wiley and Sons.

Chance, J. E., Turner, A. L., Goldstein, A. G. (1982). Development of differential recognition for own and other-race faces. The Journal of Psychology.

Dailey, M. N., Cottrell, G. W. and Busey, T. A. (1999) Facial memory is kernel density estimation (almost). Advances in Neural Information Processing Systems 11, pp. 24-30. MIT Press, Cambridge, MA.

Feingold, C. A. (1914).The influence on environment on identification for persons and things, Journal of Criminal Law and Police Science, 5, 39-51.

Jolliffe. I. T. (1986). Principal Component Analysis. Springer-Verlag.

Levin D. T. (1996). Classifying faces by race: The structure of face categories. Journal of Experimental Psychology: Learning, Memory and Cognition, 22.

Levin, D. T. (2000). Race as a visual feature: using visual search and perceptual discrimination asks to understand face categories and the cross-race recognition deficit, Journal of Experimental Psychology: General 2000 vol. 129, No. 4.

MacKay, D (2003). Information Theory, Inference, and Learning Algorithms.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. Journal of Experimental Psychology: General, 115.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. Journal of Experimental Psychology: Learning, Memory and Cognition, 14.

O'Toole, A. J.; Deffenbacher, K. A.; Abdi, H.; Bartlett, J. C. (1991) Simulation of 'other-race effect' as a problem in perceptual learning, Connection Science vol. 3, No. 2.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., Abdi, H. (1994). Structural Aspects Of Face Recognition and the Other-race Effect, Memory & Cognition, 22(2).

Phillips, P., Moon, H., Rauss, P., Rizvi S. (1997). Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition.

Sanger, T. (1989). An optimality principle for unsupervised learning. In Touretzky, D., editor, Advances in Neural Information Processing Systems, volume 1, San Mateo. Morgan Kaufmann.

Shannon, C. (1948). A Mathematical Theory of Communication, Bell System Technical Journal, Vol. 27.

Shepard, J. (1981). Social factors in face recognition. In G. Davies, H. Ellis, & J. Sbepard (Eds.), Perceiving and remembering faces. London: Academic Press.

Shapiro, P. N. & Penod, S. D. (1986). Meta-analysis of face identification studies. Psychological Bulletin, 100.

Tipping, M. E.; Bishop, C. M.(1999) Probabilistic Principal Component Analysis. Journal of the Royal Statistical Society.

Triesman A., Gormican S. (1988). Feature analysis in early vision: evidence from search asymmetries. Psychological Review 95.

Turk, M. A. & Pentland, A. P. (1991). Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3(1).

Valentine, T., Chiroro, P., & Dixon, R. (1995). An account of the own-race bias and the contact hypothesis based on a "face space" model of face recognition. In T. Valentine, Cognitive and Computational Aspects of Face Recognition: Explorations of Face Space. London: Routledge.