Unsupervised Learning of Religious Facial Features

Christopher E. Mertin

School of Computing, University of Utah

Abstract

A paper published by N.O. Rule, *et. al*, explored the possibility of humans being able to discern if someone was part of a relgious group or not [1], and was able to achieve 55% accuracy. This paper explores the use of unsupervised learning techniques and eigenfaces to perform the same task, with clustering algorithms obtaining up to 59.3% labeling accuracy on the clusters, and eigenfaces obtaining upwards of 70% accuracy on unseen data.

Transforming the Data

The source of the data came from Tinder using Pynder [2]. Faces from various parts of the US were used, namely: Seattle, Miami, Los Angeles, New York, Denver, Dallas, Las Vegas, Raleigh, and Provo. As most of the faces are not aligned, they had to be transformed to the same space.

$$\begin{pmatrix} x^\star \\ y^\star \end{pmatrix} = \begin{pmatrix} \phi_x \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \phi_y \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix}$$

Figure : Cropped Face

Figure : Transformed Face

We also need the "average face" of the two groups, done by averaging over the total number of images and summing them, resuling in

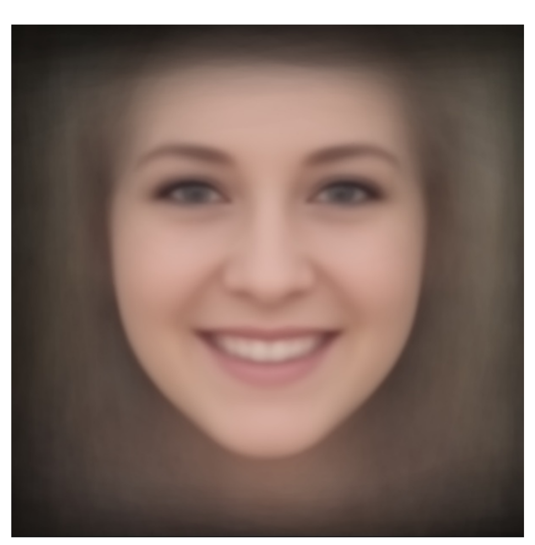

Figure : Average Mormon Face

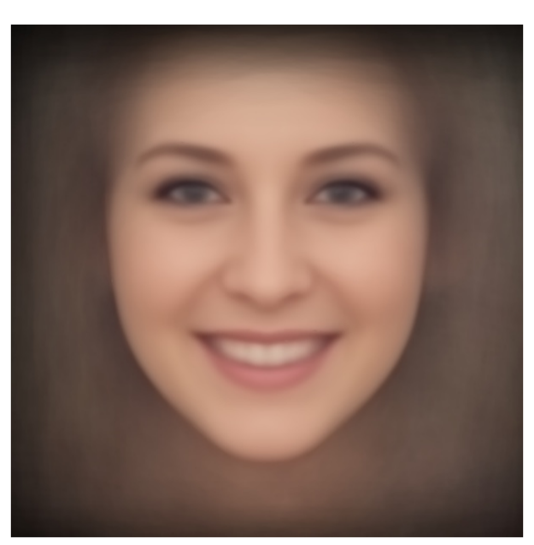

Figure : Average Non-Mormon Face

Eigenfaces

Eigenfaces uses Principal Component Analysis (PCA) for identification. It is usually used for individuals, though in this instance $\sim$ 250 Mormon faces were used. In order to determine the appropriate value of k, we needed to find the variance of the singular values

$$\frac{\sigma_i}{\sum_{i=1}^{N} \sigma_i}$$

Weights $\omega = U^T \cdot (X - \mu)$ are used for identification and can recreate the faces for various values of k.

Figure : Reconstruction of Face for various values of k

The results of testing the eigenfaces accuracy on *unseen faces* show that $k \sim 50$ gives the best results with the Cosine distance. This is the same result obtained from looking at the singular value variance.

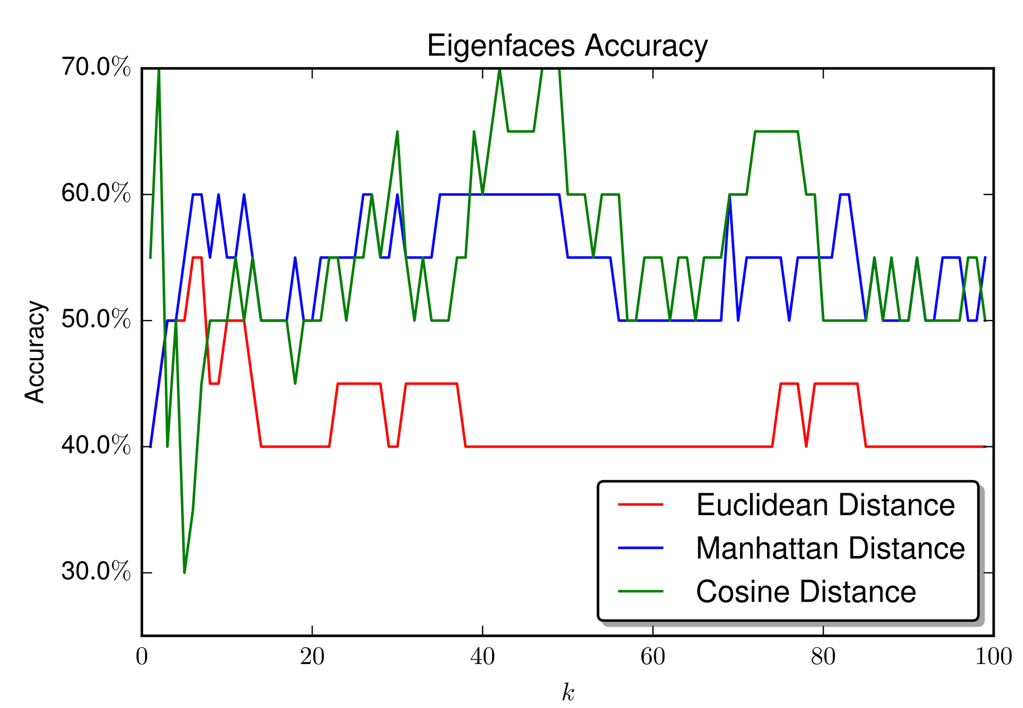

Figure : Accuracy of Eigenfaces with various metrics

Agglomerative Clustering

In order to cluster, SIFT Features were extracted and trained on. After mapping each face to the same space and find the "average face," we can look for differences in the average for dimensionality reduction.

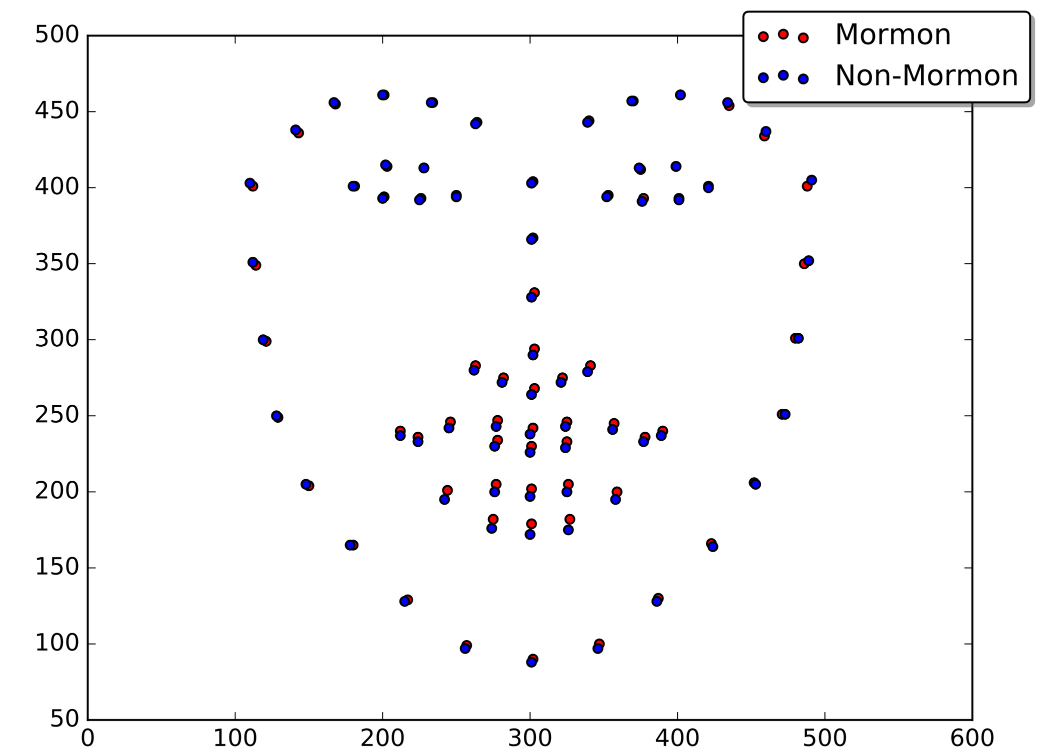

Figure : SIFT Features of "average faces"

- Hierarchial Clustering: Complete-Link with Euclidean Distance, 58.4% Accuracy

KMeans was run 1000 times on the data set to obtain the "best" cluster, and to obtain the distribution of minimums. There are two saddle points in the data which settled at 42.3% accuracy 40% of the time, and 59.3% accuracy for the remaining 60%.

After clustering, the "best" cluster was used to get the labels, and the values chosen to look at were the left eigenvectors multiplied by the eigenvalues to obtain the first principal component, $U^T\Sigma$`[0:N,0]`

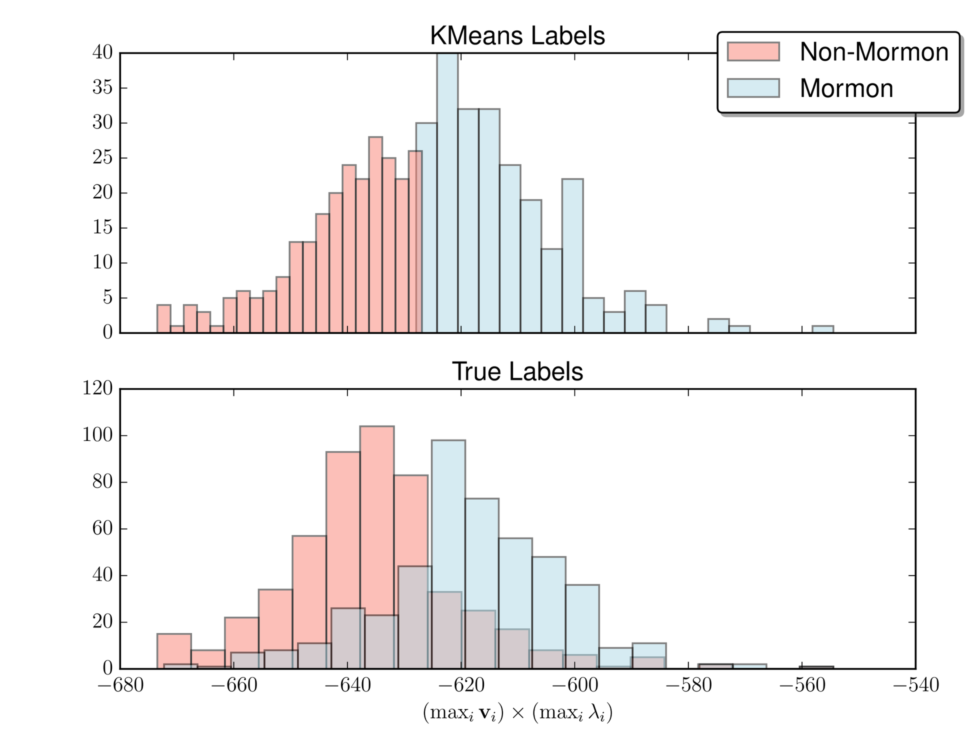

Figure : Values based on largest eigenvector and eigenvalue

Conclusion

In conclusion, the unsupervised algorithms such as hierarchical clustering and KMeans did better than humans in [1], with hierarchical clustering obtaining 58.4% accuracy and KMeans 59.3%.

The Eigenfaces algorithm performed much better with various values of k and the metric. The best metric was the Manhattan Distance which was able to achieve up to 70% labeling accuracy for $k \sim 50$. On average, the manhattan distance performed much better than either of the aforementioned clustering algorithms.

All implementations performed better than humans in [1].

References

[1] Rule, Nicholas O. AND Garrett, James V. AND Ambady, Nalini.
On the Perception of Religious Group Membership from Faces.
PLOS ONE, 5(12):1–10, 12 2010.

[2] Charlie Wolf.
Pynder: Python Client for Tinder API.
`https://github.com/charliewolf/pynder`.

[3] Tom Schad.
Cody Hoffman stared as rare non-Mormon at BYU before joining Redskins.
`http://www.washingtontimes.com/news/2014/aug/12/cody-hoffman-starred-rare-non-mormon-byu-joining-r/`.

[4] Turk, Matthew and Pentland, Alex.
Eigenfaces for Recognition.
J. Cognitive Neuroscience, 3(1):71–86, January 1991.

[5] David G. Lowe.
Distinctive Image Features from Scale-Invariant Keypoints.
Int. J. Comput. Vision, 60(2):91–110, November 2004.

Contact Information

- Web: `https://cmertin.github.io/`
- Email: `cmertin@cs.utah.edu`