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 80% accuracy on unseen data.

Transforming the Data

Need to transform the faces such that the corners of the eyes are in the same spot for each picture. This is done with the following equation.

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

This transformation applied to a cropped image can be seen below.

Figure 1: Cropped Face

Figure 2: Transformed Face

We also need the "average face" of the two groups. This can be seen in the two figures below

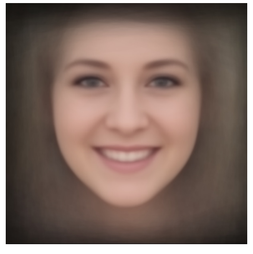

Figure 3: Average Mormon Face

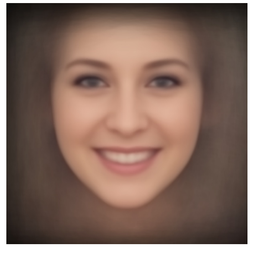

Figure 4: Average Non-Mormon Face

Eigenfaces

To understand how much influence each singular value has, we can calculate the singular value variance by

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

The variance shows that the majority of the data/images are represented within the first 50 singular values.

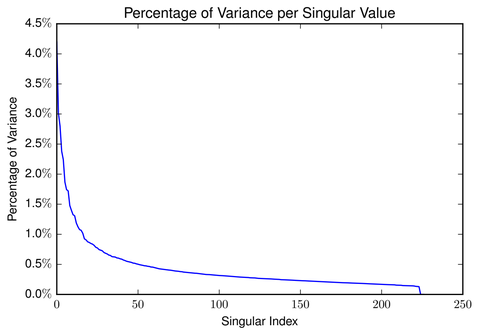

Figure 5: Accuracy of Eigenfaces with various metrics

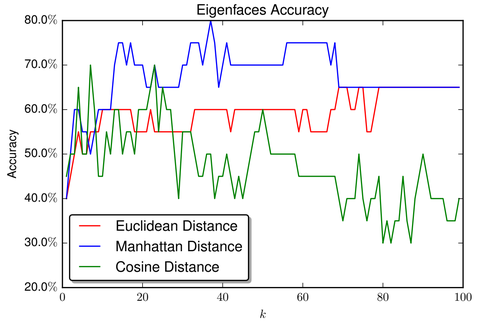

Figure 6: Accuracy of Eigenfaces with various metrics

Agglomerative Clustering

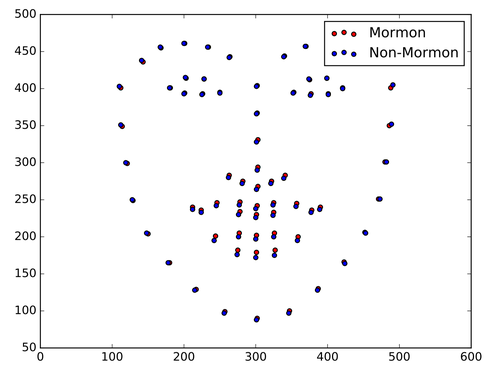

Figure 7: SIFT Features of "average faces"

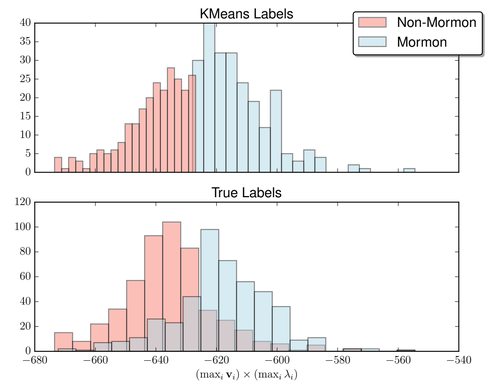

Figure 8: Labeling based on largest eigenvector and eigenvalue

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 80% labeling accuracy for $k \sim 35$. 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