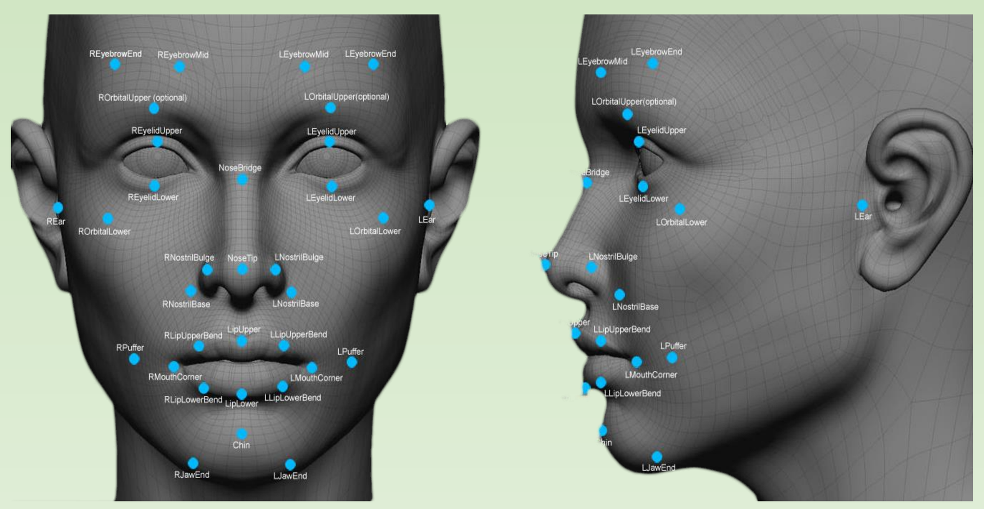
Speeding Up Image Recognition using Parallel Processing

Jian Liu*, Dr. Enyue Lu*, Dr. Yuanwei Jin +

*Department of Mathematics & Computer Science, Salisbury University, Salisbury, MD
+Department of Engineering and Aviation Sciences, University of Maryland Eastern Shore, Princess Anne, MD

Abstract

In image recognition, we represent each image as a matrix and then apply image processing algorithms to it and get useful information. In the image processing algorithms, the images are first switched to gray scale mode then we use principal component analysis (PCA) and support vector machine (SVM) to recognize the face and calculate the accuracy of classification. In this project, Jacobi Method parallel processing is used to speed up PCA in the image recognition algorithms.



Steps of Image Processing Algorithm

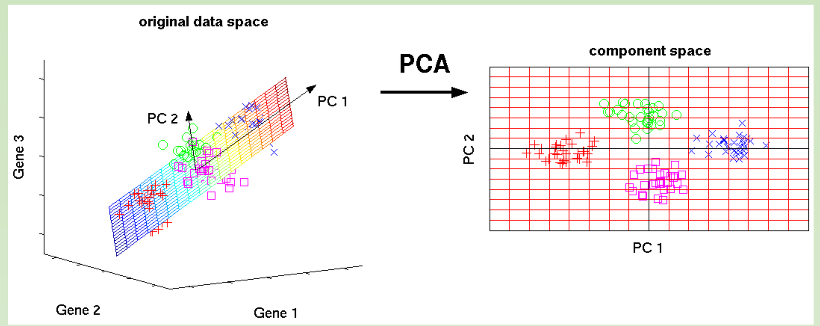
1. Extract the training sample data set
2. Modify PCA by Jacobi Method
3. Use regular PCA and modified PCA to reduce the dimension and remove the noise of the data.
4. Use SVM to train the data and get the classifier. In this step, use different parameters to make different classifiers.
5. Extract the test sample, then use regular PCA and modified PCA to reduce dimension and remove noise.
6. Use classifier obtained from step 4 to test the accuracy

FaceScrub Dataset

The dataset we used for image processing is FaceScrub. It contains 530 people, 265 females, 265 males. Each person has 200 images. WE use 100 images for training and another 100 for testing. All the images are normalized.

Principal Component Analysis

Principal component analysis (PCA) of a data matrix extracts the dominant patterns in the matrix in terms of a complementary set of score and loading plots. It is the responsibility of the data analyst to formulate the scientific issue at hand for PC projections, regressions, etc. The number of distinct principal components is equal to the smaller of the number of original variables or the number of observations minus one. This transformation is defined in such a way that the first principal component has the largest possible, and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors are an uncorrelated orthogonal basis set. PCA is sensitive to the relative scaling of the original variables.



Jacobi Transformation

The Jacobi method consists of a sequence of orthogonal similarity transformations of the form of equation.

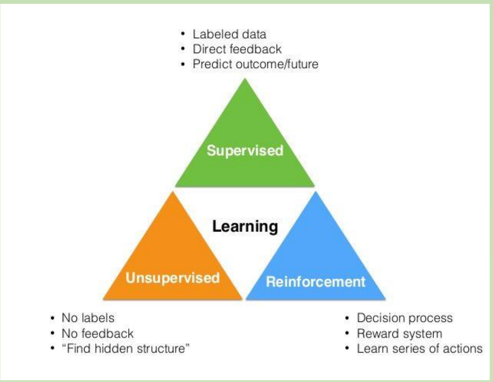
$$A \mapsto Q_{k\ell}^T A Q_{k\ell} = A'.$$

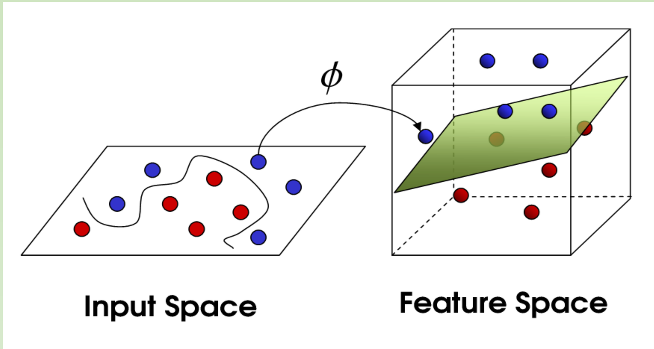
$$\begin{bmatrix} * & & & \cdots & & & * \\ & \ddots & & & & & \\ & & a_{kk} & \cdots & a_{k\ell} & & \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ & & a_{\ell k} & \cdots & a_{\ell\ell} & & \\ & & & & & \ddots & \\ * & & & \cdots & & & * \end{bmatrix} \to \begin{bmatrix} * & & & \cdots & & & * \\ & \ddots & & & & & \\ & & a'_{kk} & \cdots & 0 & & \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ & & 0 & \cdots & a'_{\ell\ell} & & \\ & & & & & \ddots & \\ * & & & \cdots & & & * \end{bmatrix}.$$

Each transformation (a Jacobi rotation) is just a plane rotation designed to annihilate one of the off-diagonal matrix elements.

Support Vector Machines

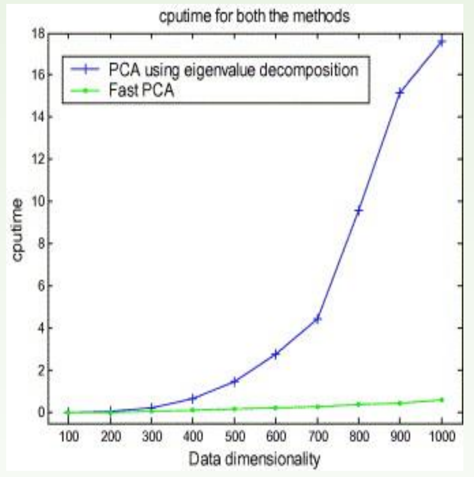
Support Vector Machines (SVM) are a set of related methods for supervised learning, applicable to both classification and regression problems. The SVM was developed by Vapnik to constructively implement principles from statistical learning theory. In the statistical learning framework, learning means to estimate a function from a set of examples (the training sets). To do this, a learning machine must choose one function from a given set of functions, which minimizes a certain risk that the estimated function is different from the actual function. Instead of computing in high dimension, we can use this function to compute in a lower dimension.





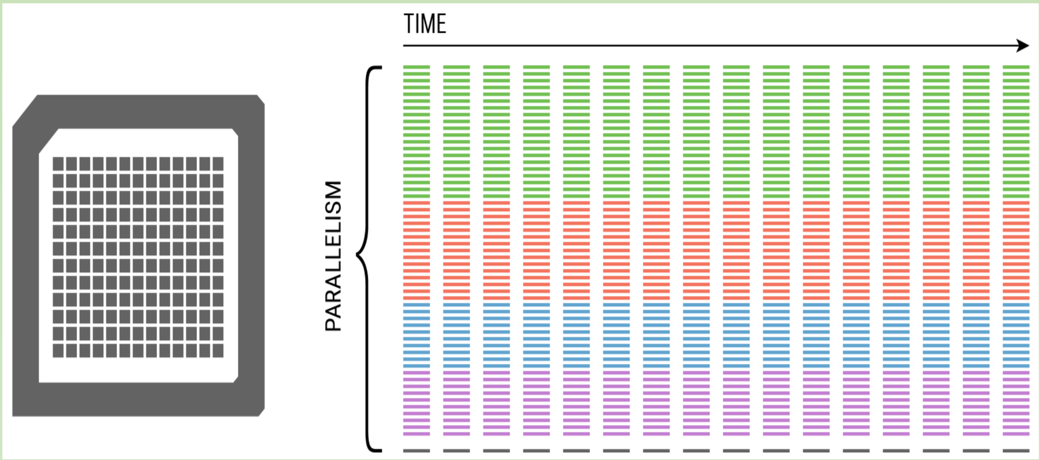
Results

By the result of the changing of the parameters several times, we can see the accuracy range from 0 to 83.3%. The PCA which is using eigenvalue decomposition is slower than fast PCA. Modifying the diagonalization part of PCA by using Jacobi Rotation in GPU is not completed. We predict the efficient will be higher and the job will be done in the future. The method needs to rebuild all the functions in the data preprocessing part.



Use Graphics Processing Unit to Compute

The graphics processing unit (GPU) has become an integral part of today's mainstream computing systems. In the diagonalization part of the PCA, we can use parallel computing to make Jacobi Method to calculate in GPU.



Conclusion and Future Work

The PCA is an efficient method to reduce the dimension of data. With the amount of data becomes larger, the advantage of using the PCA to reduce the characters in the image matrix becomes important. With reducing the dimension from n to an acceptable number k, the task of computing is less. For Support Vector Machine, with the parameter changing, the accuracy does not range higher than 83%. Picking another classifier should upgrade the accuracy of training. In the future, the data preprocessing part will be finished by using a new Jacobi Rotation Method and the classifier will be created by using Neural Network training.

Acknowledgement

Funding for this project was provided through the NSF grant CNS-1757017.
Dr. Lu for running the Salisbury REU program