Journal of Advances in Computer Research
Quarterly pISSN: 2345-606x eISSN: 2345-6078
Sari Branch, Islamic Azad University, Sari, I.R.Iran
(Vol. 10, No. 4, November 2019), Pages: 37-49
www.jacr.iausari.ac.ir

# Face Recognition Using Sparse Representations and P-Laplacian

**Homayun Motameni**✉
*Department of Software Engineering, Sari Branch, Islamic Azad University, Sari, Iran*
motameni@iausari.ac.ir

**Received: 2019/01/17; Accepted: 2019/05/01**

**Abstract**

*Face recognition is one of the most important identification tools in biometrics. Nowadays, the topic of face recognition has many applications in various fields, including public security, identity identification, protection of important and sensitive places, access control, video surveillance, and so on. Two important issues in face recognition applications are speed and accuracy in detection. Various studies have shown that face recognition through Sparse Representation Classification (SRC) works very well. The purpose of this paper is to propose a fast and efficient method for the sparse representation-based face recognition .Due to the fact that retrieving the Sparse Representation based on L1 norm optimization for a large dictionary has a large computational volume, a Smooth L0 norm optimization (SL0) method is used. Also, due to the fact that one of the challenges of face recognition is the existence of brightness changes in images, so we use the P-Laplacein algorithm in the feature extraction step to give us more complete information about the face image by recognizing the edge. As the simulation results on the Extended Yale B and AR database show, the proposed hybrid method has a higher detection rate than the sparse display method.*

***Keywords:*** *Face Recognition, Sparse Representation, Smoothed L0-Norm, L1-Norm, Laplacian Face, P-Laplacian*

## 1. Introduction

Face recognition is one of the most important tools of identification in biometrics. Nowadays, face detection is a common practice in many areas, including public security, identification, protection of critical and sensitive locations, access control, video monitoring, and so on. In addition, face detection systems are more prone to use in new areas such as human-computer-related relationships, Internet services such as Internet shopping and more. On the other hand, with the advancement of technology, we are witnessing the increasing growth of human-computer interactions, in which the issue of face recognition is one of the basic steps. Various algorithms for face recognition have been proposed. Current research on face recognition usually involves feature extraction and classification selection. In feature-based face recognition algorithms, features that effectively display face images can increase the recognition rate and reduce the cost of calculation and storage. However, since real-world face images are mostly nonlinear and experience significant changes in brightness, facial expressions, and gestures, feature-based extraction algorithms cannot be used for Full control of various changes was employed. Therefore, the classifier-based algorithm, which is another important algorithm in the field of face recognition algorithms, should be used. These algorithms are common and well-known algorithms that are widely used in face recognition. Two important issues that are important in face recognition applications are speed and accuracy in diagnosis. Over the past decade, effective efforts have been made to improve the performance of detection speed and accuracy [1]. In this article, considering the importance of face recognition and its applications, a method has been proposed that is effective in recognizing faces with high accuracy and speed .Wright et al., In order to obtain the minimum reconstruction error in classification-based algorithms such as NN and NS, proposed the Spars Representation method (SRC) [2], which, in the presence of relative obstruction in facial images, performed better than Provides NN and NS. the experimental image is modeled as a linear combination with sparce coefficients of all training images. Next, reconstruction errors will be obtained between the experimental and representation images in each training data set. Finally, the test image is assigned to the batch label with the least residual residual errors. The test results show that SRC will perform very well, especially in the field of obstruction and tolerance of noisy environments.

Here, the Sparce representation method is used to classify the extracted features. This method has a better performance than other classification methods. One of the most important and distinguishing features of this method from other methods is that this method is resistant to dimming of a part of the face. This method uses $L_1$–norm optimization or linear programming (LP) or basic search (BP) to find the thin answer. Then, to minimize the computational complexity and execution time of the program, the smooth norm $L_0$ (SL0) [3] method, which has a higher convergence speed than BP, was used instead of the BP algorithm. SL0 algorithm In case of unique conditions, this algorithm is superior to BP algorithm in terms of both performance and speed [3].

One of the most common practices in image analysis is edge detection because the edge is the boundary between an object and its background. The image occurs. The greater the change in level, the easier it is to detect the edge. In fact, points in the image that have sudden brightness changes are often called edges or edge points. Edge points usually include object borders and other types of brightness changes, as well as noise edges[4].

There are three steps in edge detection that include:

- Edge recognition patterns include three stages: filtering, derivation, and detection.
- In the filtering step, the image is passed through a filter to remove noise.
- In the derivation stage, situations in the image are highlighted by changes in noise intensity.

- Finally, at the detection stage, the edge points are determined by thresholding the points highlighted in the derivation stage.

One of the edge detection algorithms is Laplacein algorithm.

In this paper, the P-Laplacian method can be used in the preprocessing stage.

## 2. Brief Review of Laplacian

Laplacianface is a recently proposed linear method for face representation and recognition. The Laplacianfaces method preserves the local structure of the image space. This method is linear and is computationally more efficient than other nonlinear techniques. Uses an appearance-based approach to recognizing the human face. This method reduces the size of the face image. It actually creates a subspace that explicitly considers the structure of the facial manifold. The structure of the face manifold is modeled by the nearest neighbor diagram that uses the local structure of the image space[5].

Given a set of face images $\{x_1,...,x_n\} \subset R^m$, let $X = \{x_1,x_2...,x_n\}$ .Let S be a similarity matrix defined on the data points. Laplacianface can be obtained by solving the following minimization problem:

$$a_{opt} = \arg\min_{a} \sum_{i=1}^{n}\sum_{j=1}^{n}\left(a^T x_i \text{-} a^T x_j\right)^2 S_{ij} = \arg\min_{a} a^T XLX^T a \qquad (1)$$

with the constraint

$$a^T XDX^T a=1 \qquad (2)$$

where L = D - S is the graph Laplacian [4] and $D_{ii} = \sum_j S_{ij}.D_{ii}$ measures the local density around $x_i$. Laplacianface constructs the similarity matrix S as shown in the equation at the bottom of the next page. Here, $S_{ij}$ is actually heat kernel weight, the justification for such choice and the setting of the parameter t can be referred to [3]. The objective function in Laplacianface incurs a heavy penalty if neighboring points $x_i$ and $x_j$ are mapped far apart. Therefore, minimizing it is an attempt to ensure that if $x_i$ and $x_j$ are “close,” then $y_i = \left(=a^T x_i\right)$ and $y_j = \left(=a^T x_j\right)$ are close, as well [9]. Finally, the basis functions of Laplacianface are the eigenvectors associated with the smallest eigenvalues of the following generalized eigen-problem:

$$XLX^T a=XXDX^T a \qquad (3)$$

$XDX^T$ is nonsingular after some preprocessing steps on X in Laplacianface; thus, the basis functions of Laplacianface can also be regarded as the eigenvectors of the matrix $(XDX^T)^{-1}XLX^T$ associated with the smallest eigenvalues. Since $(XDX^T)^{-1}XLX^T$ is not symmetric in general, the basis functions of Laplacianface are nonorthogonal. Once the eigenvectors are computed, let $A_k = [a_1; \ldots; a_k]$ be the transformation matrix. Thus, the Euclidean distance between two data points in the reduced space can be computed as follows:

$$\begin{aligned} \text{dist}\left(y_i,y_j\right) &= \left\|y_i - y_j\right\| \\ &= \left\|A^T x_i - A^T x_j\right\| \\ &= \left\|A^T (x_i - x_j)\right\| \\ &= \sqrt{\left(x_i - x_j\right)^T AA^T \left(x_i - x_j\right)} \end{aligned} \quad (4)$$

If A is an orthogonal matrix, $AA^T = I$ and the metric structure is preserved.

***2.1 P-Laplacian***

Given a set of similarity measurements, the data can be represented as a weighted, undirected graph G=(V,E), where the vertices inVdenote the data points and positive edge weights in W encode the similarity of pairwise data points[6]. We denote the degree of node $i \in V$ by $d_i = \sum_j w_{ij}$. Given function $f:V \rightarrow R$, the p-Laplacian operatoris defined as follows:

$$\left(\Delta_p^w f\right)_i = \sum_j w_{ij} \phi_p \left(f_i - f_j\right) \quad (5)$$

Where $\phi_p\left(x\right) = \left|x\right|^{p-1} \text{sign}\left(x\right)$ .Note that $\varphi_2(x)=x$, which becomes the standard graph Laplacian. In general, the p-Laplacian is a nonlinear operator. The eigenvector of p-Laplacian is defined as following:

**Definition:** $f:V \rightarrow R$ is an eigenvector of p-Laplacian $\Delta_p^w$ , if there exists a real number $\lambda$, such that

$$\left(\Delta_p^w f\right)_i = \lambda \phi_p \left(f_i\right), \quad i \in V \quad (6)$$

$\lambda$ is called as eigenvalue of $\Delta_p^w$ associated with eigenvector f. One can easily verify that when p=2, the operator $\Delta_p^w$ becomes the regular graph Laplacian $\Delta_2^w = L = D - W$ ,where D is a diagonal matrix with $D_{ii} = d_i$,and the eigen-vectors of $\Delta_p^w$ become the eigenvectors of L. The eigenvector of p-Laplacian is also called.

### 3. Classification based on Sparce Reprentation

The fundamental issue in identifying an object is to determine the test data class using labeled training samples of k distinct classes. ni training samples of i-th class are placed in columns of the matrix $A_i=\left[V_{i,1},V_{i,2},....,V_{i,n}\right]\times R^{m\times n_i}$ . In the face recognition, a black and white image with a dimension of $w\times h$ is represented through the vector R by sequentially placing the image columns. The columns of Ai denote the training facial images related to i-th person.

A variety of models and representations have been presented for $A_i$. According to one of the simplest and most effective models, every image or feature extracted from a class belongs to a linear subspace. Based on another model called as "face subspace", the same facial images under lightness changes cover the approximate subspace with a low dimension in space $R^m$ [7, 8], i.e. each image class occupies a small part of the m-dimensional space. For simplicity, it is assumed that the training samples of each class form a subspace. Hence, each test vector sample $y\in R^m$ of i-th class can be approximated as a linear combination of the training sample vectors of i-th class with scalar values $a_{i,j}\in R$, $j=1,....,m$:

$$y=a_{i,1}v_{i,1}+a_{i,2}v_{i,2}+...+a_{i,n_i}v_{i,n_i} \qquad (7)$$

Since the test sample belonging to i-th class is not clear, we define the matrix A by putting in sequential order the matrices $A_i$ as follows:

$$A=\left[A_1,A_2,...,A_k\right]\in R^{m\times n} \qquad (8)$$

Therefore, the linear representation of the test vector sample $y\in R^m$ is rewritten as follows:

$$y = Ax_0\in R^m \qquad (9)$$

Where $x_0=\left[0,...,0,a_{i,1},...,a_{i,n},0,...,0\right]^T$ is the coefficients vector. All entries of the coefficients vector are zero except for i-th class. Since entries of the vector $x_0$ specify the identity of the test sample y, the solution of the linear equation $y=Ax$ is important. Compared to local NS and NN methods, the methodology used here is global. It is clear that the solution of the linear equation $y=Ax$ depends on the ratio of the number of equations m to the number of unknowns n. if $m>n$ , then the linear equation $y=Ax$ will be over determined and the unique solution can be found. For face recognition, the equation $y=Ax$ is usually underdetermined (m<n) and then the solution is not unique. To solve this problem, we can usually use the $L_2$-norm solution as follows:

$$(L_2) \qquad \hat{x}_0 = \operatorname{argmin} \|x\|_2 \ \text{s.t}, y = Ax \qquad (10)$$

Although the optimization of this problem can be simply solved by quasi-inverse matrix A, but because the solution is not sparse (in general), this solution does not contain specific information for the face recognition application. Considering the nature of the problem and concept of the face subspace model, SR techniques are suggested to solve this problem. Each very sparse solution can identify the identity of the test sample Y according to the stated concepts. Hence, the method for finding the sparsest solution of Equation $y=Ax$ by solving the following optimization problem is proposed.

$$(l_0) \qquad \hat{x}_0 = \operatorname{argmin} \|x\|_0 \qquad \text{s.t}, y = Ax \qquad (11)$$

$\|\cdot\|_0$ denotes the L0-norm and counts the number of non-zero entries. It is shown that if the matrix A is random and the equation $y=Ax$ has the answer that the number of its non-zero elements is less than half the number of equations $m/2$, this answer is unique, i.e. $\hat{x}_0 = x_0$ [9].

The general theorem of uniqueness of sparse solution is expressed by the definition of the spark of the matrix A. The problem of finding the sparsest answer is a hard problem because $L_0$-norm is not derivable. Several approximate approaches have been presented to solve this problem. In the following, we summarize the two methods proposed to solve this problem.

### *3.1 Finding the sparse solution by L1-norm minimization*

Many methods of solving (11) using the cost function, which is a measure of response sparseness, try to find the sparsest answer by solving an optimization problem. If the selected cost function is a better criterion for being sparse the vector x, then the accuracy of the final solution is higher. On the other hand, the desirable cost function is a function that simplifies the solving of the optimization problem.

Some methods use the $L_1$-norm minimization metric to measure the vector sparsity, namely:

$$(L_1) \qquad \hat{x}_0 = \operatorname{argmin} \|x\|_1 \qquad \text{s.t}, y = Ax \qquad (12)$$

Substituting of $L_0$-norm with $L_1$-norm converts the optimization problem into a convex one, which has appropriate methods (e.g. BP). In [10], it has been proven that in majority of underdetermined linear systems, $L_0$-norm and $L_1$-norm minimization leads to one solution, which is the sparsest one. Of course, in this method, the uniqueness condition is very limited compared to that of the answer (11). This is one of the reasons that some researchers have focused on ideas that directly solve the $L_0$-norm problem.

In [2], researchers utilize the noise model (12) for face recognition as follows:

$$(L_1) \qquad \hat{x}_0 = \text{argmin}\|x\|_1 \quad \text{s.t,} \|y-Ax\|_2 \le \varepsilon \qquad (13)$$

Here we use the model $y=Ax_0+z$ to represent the test sample, which $z \in R^m$ includes the noise model with limited energy $\|z\|_2 \le \varepsilon$.

### *3.2 Finding the sparse solution by smoothed L0-norm minimization*

As shown below, the main idea of this method is to use a continuous and smooth function for the $L_0$-norm approximation of the vector x:

$$\|x\|_0 \approx m-\sum_i \exp\left(\frac{-x_i^2}{2\sigma^2}\right) = m-F_\sigma(x) \qquad (14)$$

If σ tends to zero, the approximation sign is converted to equality. Therefore, maximizing the function $F_\sigma(x)$ for a small σ is equivalent to minimizing the $L_0$-norm. Therefore, the optimization problem will be as follows:

$$(SL0) \qquad \hat{x}_0 = \text{argmax} F_\sigma(x) \qquad \text{s.t,} y=Ax \qquad (15)$$

To find the answer to the above problem, we can use the steepest descent maximization algorithm. For details of the optimization algorithm of problem (15) called as SLO, refer to [3]. An important feature of this algorithm is the high convergence rate and its good performance compared with the BP algorithm.

### *3.3 Classification based on sparse representation*

For the new test sample y associated with one of training set classes, firstly we obtain SR by using (12) or (13) or (14). Ideally, non-zero entries in $x_0$ estimation are related to i-th class columns. Therefore, we can easily assign the test sample y to that class. However, this is not the case due to model error and noise. In fact, small non-zero entries corresponding to other classes can also be found. Accordingly, many classifications can be designed for the identification area. In one of the easiest ways, the class with the largest non-zero entry is considered as the winner. Nevertheless, the method presented in [2] is a more logical and subjective one. This approach examines which linear combination of classes training samples with coefficients obtained from SR offers a better approximation of the test sample y.

For each class i, vector $\delta_i(\hat{x}_0) \in R^m$ derived from $\hat{x}_0$ is a vector in which entries corresponding to classes other than i are set to zero. By this definition, the approximation of the test sample y using the training samples of i-th class is equal to $y_i = A\delta_i(\hat{x}_0)$ . So the classification of y is based on the best approximation, that is, the class with the least estimation error is the winner one:

$$\min_i r_i(y) = \min_i \|y-A\delta_i(\hat{x}_0)\|_2 \qquad (16)$$

The classification algorithm is summarized in the following.

SRC algorithm's summary:

1.Input: matrix of training samples $A \in R^{n\times m}$ and test sample $y \in R^{m}$.

2.normalizing the columns of matrix A.

3.obtaining SR for the test sample using (11) or (12) or (13).

4.calculating the estimation error for all classes by $r_i(y)=\|y\text{-}A\delta_i(\hat{x}_0)\|_2$ ,i=1,...,k .

5.output: identifying y with $\min_i r_i(y)$.

Figure 1 briefly illustrates the steps involved in this algorithm. In this example, the down-sampled image is used as a feature, and the test image belongs to the first class. As shown in SR, the coefficients related to class 1 have a larger size. In this figure, images associated with two larger coefficients have also been shown in SR. These images correctly identified the person's face. The diagram of $r_i(y)$ also confirms this $r_1(y)$. has the lowest value and assigns the test sample to class 1.

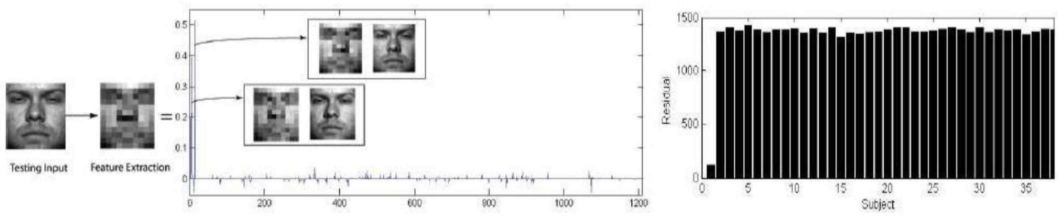


***Figure 1: structural view of SRC algorithm [5]***

## 4. Proposed Approach

In this paper, the Sparce Representation method is used to classify the extracted features. This method has a better performance than other classification methods. One of the most important and distinguishing features of this method from other methods is its resistance to dimming of a part of the face. In this method, soft optimization method $L_1$ or linear programming (LP) or basic search (BP) is used to find the thin answer. Due to the fact that retrieving the Sparce representaion based on $L_1$ soft optimization for a large-sized dictionary has a large computational volume, in order to minimize the computational complexity and execution time of the program, the smooth $L_0$-norm method (SL0) [3], which has Convergence speed is higher than BP, used instead of BP algorithm. The SL0 algorithm is superior to the BP algorithm in both performance and speed when the unique conditions are properly observed [3].

We have also used the P-Laplacein method in the feature extraction step to increase the recognition accuracy in face recognition and face verification under variable brightness and to find the edge points. Edge points usually include object borders and other types of brightness changes as well as noisy edges. The simulation results on the Extended Yale B database show that the SL0 method together with the P-Laplacein method has a high accuracy in identifying faces with less computational volume and higher speed than the $L_1$ norm based method.

### *4.1 Combination of SRC and P-Laplacian algorithms*

In images, the border between two areas that has a significant difference in brightness, color, or texture is usually called the edge. One of the advantages of edge extraction in the image is the ability to separate and distinguish objects from the background. This special edge advantage has led to the edge recognition process being used in a variety of applications, including zoning, extraction of boundary features, and shape description.

Dividing an image into objects and backgrounds by following the delimiter boundary between them is an important step in analyzing and interpreting images. In this paper, a new method for detecting Laplacian-based edge called P-Laplacian, which is a more general form of Laplacian, is presented. As in the previous section, the thin display method is used to classify the extracted features. We also propose the P-Laplacian method for feature extraction in this section. Figure 2 describes the block diagram of the new face recognition system.

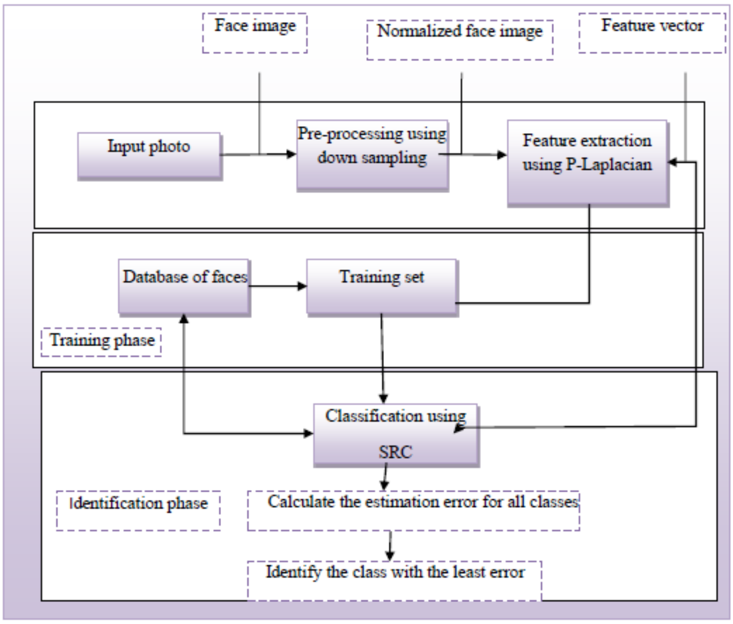


*Figure 2. Block diagram of the new face recognition system*

## 5. Database

In this article, two databases in this field have been used to evaluate the performance of the proposed algorithm. Here, we will first introduce these databases.

### *5.1 Extended Yale B data base*

This database contains 2414 face images from the front view of 38 people. All images in the size of 162 × 192 were taken in a variety of laboratory controlled lighting conditions.

### *5.2 AR data base*

The AR database contains more than 4,000 face images for 126 subjects, and the size of each image is 165 by 120. For each subject, the database in question has two separate sessions that are taken under different conditions such as changes in brightness, face and gesture. In this experiment, a subset containing 50 male and 50 female subjects is selected. In the first experiment, each subject had 14 face images taken only by changing the brightness and different facial expressions. We select 7 images from session 1 as training samples and another 7 images as test samples.

## 6. Simulation Results

in this section, the classification performance is examined by sparce representation with sampling and p-laplacian properties. As mentioned, this paper uses the Extended Yale B and AR databases.We have calculated the correct detection percentage of the algorithm for a number of different properties 36,54, 132 and 504. These numbers are proportional to the down sampling ratios of 1/32, 1/24, 1/16 and 1/8, respectively. The P-Laplacian algorithm was tested for p = 2, p = 2.5, p = 3 and p =3.5.

Here the performance of Sparce representation classification for BP, SL0, BP+P-Laplacian and SL0 P-Laplacian feature extraction method (SL0 + P-Laplacian) is calculated and their results are compared. The test results are shown in Figure 3.

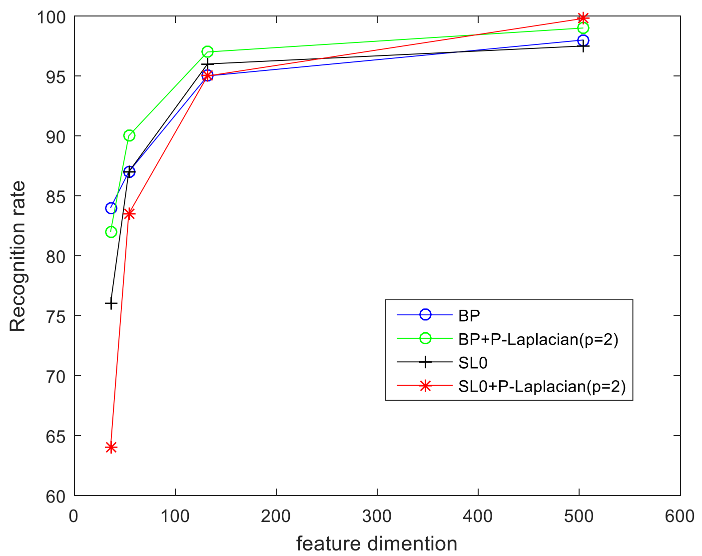


***Figure 3. Face recognition detection rate for Extended Yale B database***

As the simulation results show, the combination of P-Laplacian algorithms and sparce representation base on SL0 is a method with low computational volume and high accuracy compared to the proposed BP algorithm in [2]. In fact, the SL0 + P-Laplacian algorithm performs better than other algorithms when the number of features increases. In other words, when the number of features is large, using the SL0 + P-Laplacian algorithm has a similar function to the BP + P-Laplacian algorithm, while having a smaller computational volume and much higher speed. As shown in Figure 3, the BP and SL0 algorithms, respectively. They have an identification rate of approximately 98 and 97.5 percent. If these two algorithms, along with the P-Laplacian preprocessing operation, have reached the ideal detection rate of almost 100%.

In the following simulation, we obtain the percentage of face image recognition in the combined method of Sparse representation and P-Laplacein algorithm for different values of P. As shown in Figure 4, in the Extended Yale B database, the percentage of face recognition per p = 2.5 is better than p = 2, which is 99%. Also, the detection rate increases for p = 3 compared to p = 2.5, which is 99.8%. But the interesting thing about this simulation is that the capacity to increase p to have a higher detection rate is limited because then, by setting the value of p = 3.5, the detection rate decreases slightly compared to p = 3 and reaches 99.3 % arrives. In this experiment, the best result was obtained by placing an approximate value of p = 3.

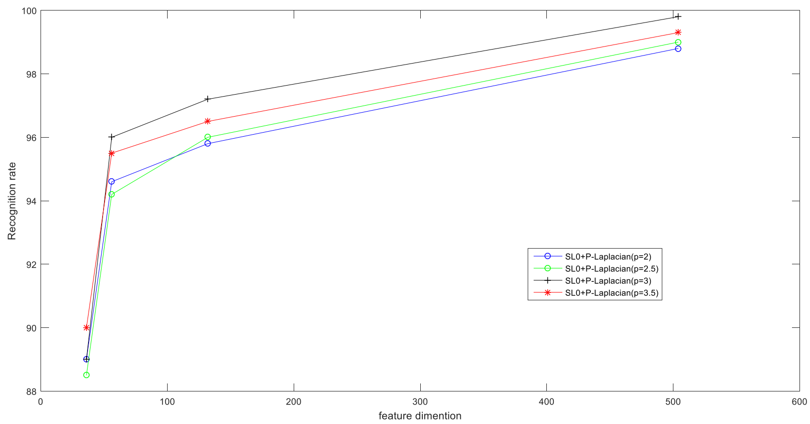


***Figure 4: Face recognition detection rate for Extended Yale B database***

In Figure 5, we see the simulation results of this proposed idea in the AR database, which shows that the proposed hybrid method in the P = 3 mode in the P-Laplacian algorithm has the highest detection rate of 99.9% compared to other p values.

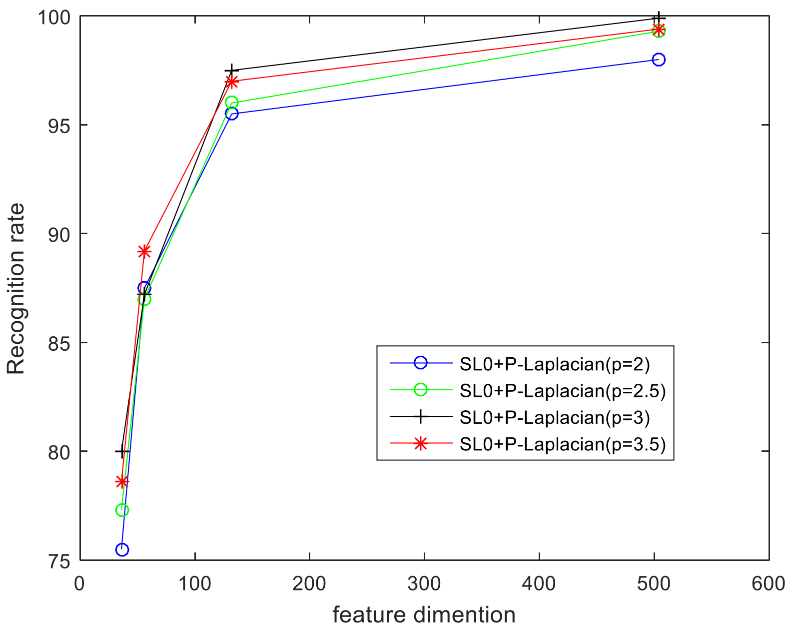


***Figure 5: Face recognition detection rate for AR database***

## 7. Conclusion

In this paper, while evaluating the performance of the P-Laplacian method in the preprocessing stage and Sparse representation Classification in face recognition, the performance of the SL0 algorithm or smooth $L_0$ norm in face recognition application was investigated. In this paper, the idea of SL0 algorithm along with P-Laplacian algorithm was used to achieve high speed and accuracy in the application of classification based on Sparce representation. The SL0 + P-Laplacian method, while having a high speed, has the same performance as the Sparce representation classifier with BP. In the case of face recognition, the conditions do not appear to be correct or the algorithm gets stuck in local minima. Therefore, to improve the performance and take advantage of the high speed of the SL0 algorithm, the idea of displaying weight and weight with the SVD algorithm can be used as future work.

## References:

[1] Xingang Liu ∗, Lingyun Lu, Zhixin Shen, Kaixuan Lu,” A novel face recognition algorithm via weighted kernel sparse Representation”,future computer.

[2] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Yi Ma, “Robust face recognition via sparse representation,” IEEE Transaction on Pattern Analysis and Machine Intelligence, , no. 2, pp. 210–227, February 2009.

[3] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, “A fast approach for overcomplete sparse decomposition based on smoothed L0 norm,” IEEE Trans. on Signal Proc, vol. 57, no. 1, pp. 289–301, January 2009.

[4] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, “Face recognition using Laplacianfaces,” IEEE Transactions on pattern.

[5] Deng Cai, Xiaofei He, Jiawei Han, Senior Member," Orthogonal Laplacianfaces for Face Recognition," IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 15, NO. 11, NOVEMBER 2006.

[6] Dijun Luo · Heng Huang · Chris Ding · Feiping Nie," On the eigenvectors of p-Laplacian," Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX,USA.

[7] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, "Face recognition using Laplacianfaces," IEEE Transactions on pattern Analysis and Machine Intelligence, vol. 27, no. 3, pp.328–340, 2005.

[8] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 2, pp. 218–233, 2003.

[9] D. Donoho and M. Elad, "Optimal sparse representation in general (nonorthogonal) dictionaries via L1 minimization,"Proceedings of the National Academy of Sciences of theUnited States of America, pp. 2197–2202, March 2003.

[10] D. Donoho " For most large underdetermined systems of linear equations the minimal L1 norm solution is also the sparset solution," Tech. Rep., 2004.