Sparse Representation Classification via Fast Matching Pursuit for Face Recognition

Michael M. Abdel-Sayed, Ahmed Khattab, Mohamed F. Abu-Elyazeed
EECE Department – Cairo University, Giza, Egypt
michaelmelek@hotmail.com, akhattab@ieee.org, mfathyae@yahoo.com

***Abstract*—Face recognition is a widely studied pattern recognition problem. One of the most crucial components of face recognition problems is classification. Sparse representation-based classification (SRC) has been recently proposed to considerably improve the classification performance by using the compressed sensing theory. However, SRC utilizes ℓ_1 minimization for recovery. Despite being optimal, ℓ_1 minimization is computationally expensive, and hence, not applicable in real-time applications. In this paper, we present the Fast Matching Pursuit (FMP) which is a compressed sensing recovery algorithm that results in a recognition time that is only 4% to 10% of that of ℓ_1 minimization and approximately half the time of existing related matching pursuit approaches. This significant speedup does not come at the expense of any degradation in the recognition rate.**

I. Introduction

A typical face recognition problem can be stated as follows: Given a still or video image of a scene, it is required to identify a person using a database of face images. Compared to other biometric features, face recognition is natural, non-intrusive and can be performed using images taken at a distance. Face recognition is widely used in many applications such as security, access control, and surveillance. The solution of the face recognition problem typically involves two processes: feature extraction and classification. Several classification techniques exist including nearest neighbor, nearest subspace, support vector machines, and neural networks [1].

Recently a classification method termed Sparse Representation-based Classification (SRC) [2], [3] has been developed based on the compressed sensing theory [4]. Compressed sensing is a sampling technique that is capable of reconstructing sparse (or compressible) signals from samples collected at a much lower rate than the Nyquist rate. The idea of SRC is based on representing the sample to be classified as a linear combination of training samples of the same subject only. This is a sparse representation that involves only a small fraction of the overall training database. In order to obtain such a sparse representation, compressed sensing recovery techniques are utilized [4]. Traditionally, ℓ_1 minimization is utilized. However, it is computationally expensive and not suitable for most real-time applications. Therefore, other algorithms have been proposed to reduce its complexity without affecting the accuracy. Among the utilized algorithms are fast ℓ_1 algorithms, such as the homotopy method [5], and greedy recovery algorithms, such as orthogonal matching pursuit (OMP) [6], which result in improvement in complexity.

In this paper, we present the Fast Matching Pursuit (FMP) algorithm that significantly outperforms both fast ℓ_1 algorithms and existing greedy algorithms such as OMP. FMP is based on two main ideas. First, it targets the selection of an optimum number of correlation values per iteration. A reduced set of such correlation values is formed, from which the top-magnitude elements are selected. The number of selected elements is adapted based on the distribution of the correlation values. Second, least square minimization is performed iteratively, avoiding large matrix inversion. In each iteration, instead of directly obtaining the inverse matrix required for minimization, it is obtained iteratively from data in the previous iteration. The inverse is updated first by adding columns corresponding to the added elements. Then, it is updated by removing columns corresponding to the pruned elements. Consequently, FMP significantly speeds up the recognition process without causing degradation in the recognition rate.

The rest of this paper is organized as follows. Section II presents an overview of the SRC. We present the FMP algorithm in Section III, and evaluate its performance in Section IV. Section V concludes the paper.

II. Sparse Representation-based Classification Overview

Sparse Representation-based Classification (SRC) was developed to exploit compressed sensing to improve the recognition rates compared to traditional classification approaches such as nearest neighbor [2]. Given a training set that includes samples from n subjects and a test sample y that is required to be classified, the test sample y can be represented by a linear combination of the training samples in the training set which are associated with the corresponding subject. If Z_i denotes the matrix whose columns are the training samples of subject i, the augmented matrix Z is constructed as:

$$Z = [Z_1, Z_2, \ldots, Z_n] \tag{1}$$

Let x be the vector that contains the weights required to form the test sample y from the training samples. Hence, y is expressed as :

$$y = Zx = [Z_1, Z_2, \ldots, Z_n]x \tag{2}$$

As illustrated in Fig. 1, the elements of x are either large-magnitude elements at indices corresponding to the subject

978-1-5386-1359-7$31.00 ©2017 IEEE

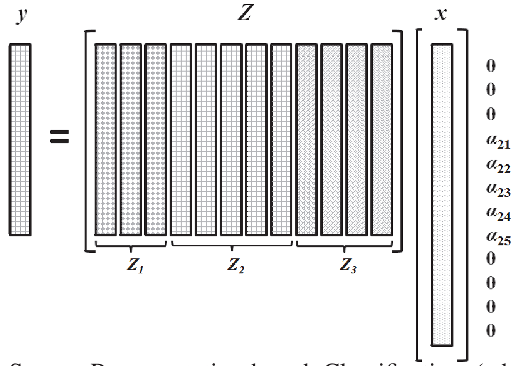

Fig. 1: Sparse Representation-based Classification (adapted from [7]).

that the test sample belongs to, or very small (or zero) elements at the remaining indices. Therefore, x is sparse vector, and hence, it can be reconstructed using ℓ_1 minimization as follows:

$$\hat{x} = \arg\min_x \|x\|_1 \text{ subject to } y = Zx \tag{3}$$

Given the reconstructed vector $\hat{x}$, the test sample y is reconstructed from linear combinations of the columns of Z using indices of $\hat{x}$ associated with each subject individually. The sample is classified to correspond to the subject with the least reconstruction error. SRC is shown to outperform traditional classification techniques, including nearest neighbor and nearest subspace [2].

III. Fast Matching Pursuit for Fast Recovery

Instead of employing ℓ_1 minimization for compressed sensing recovery as in SRC, we present the Fast Matching Pursuit (FMP) greedy recovery algorithm to significantly reduce the reconstruction time while maintaining the high recognition rate of ℓ_1 minimization. FMP is composed of four components: support identification, signal estimation, pruning, and residual calculation.

Support Identification: In compressed sensing, a measurement vector (also called the samples) $y \in \mathbb{R}^m$ is obtained by randomly sampling a sparse signal $x \in \mathbb{R}^n$ using a sensing or measurement matrix $\Phi \in \mathbb{R}^{m\times n}$, where $m \ll n$. To reconstruct the sparse signal from its samples, its non-zero indices should be *iteratively* identified first to form the support set. In each iteration, the measurement vector is correlated with the columns of the sensing matrix Φ to obtain a correlation vector g. Since the non-zero indices are expected to have high magnitude of correlation, the top magnitude elements of g are selected and their indices are merged with the identified support set. Selection in the FMP algorithm follows a selection strategy that we previously developed in [8] using the double-thresholding technique. This selection strategy adaptively selects elements from a reduced set of the correlation vector. First, the elements from which we perform selection are reduced to a set containing the top magnitude elements (using a certain threshold β). Then, elements of larger magnitude than a fraction $0 < \alpha < 1$ of the maximum element are selected from the reduced set, and their indices are added to the support set. Properly selecting α and β leads to the selection of an optimum number of elements per iteration. Using exhaustive simulation of α and β values, we found that moderate values of both $\alpha \in [0.5, 0.7]$ and $\beta \in [0.15, 0.75]$ result in the best performance because the number of selected elements per iteration becomes neither too large nor too small. Furthermore, the simulation results indicated that the algorithm performance is *not* sensitive to the specific α and β values as long as they are in the aforementioned optimum range.

Signal Estimation: A signal estimate $\hat{x}$ is formed based on the identified support set using least square minimization. $\hat{x}$ minimizes $\|y - \Phi\hat{x}\|_2$ while having nonzeros at indices from the identified support set. This is done by multiplying y by the pseudo-inverse of Φ_T, where Φ_T is a matrix containing columns of Φ at indices from the identified support set T. Typically, the pseudo-inverse of Φ_T is obtained directly as ${\Phi_T}^\dagger = (\Phi_T^T\Phi_T)^{-1}{\Phi_T}^T$, which involves matrix inversion. In contrast to prior literature including ours [8], we propose to perform signal estimation iteratively in FMP. The inverse is first updated by adding columns of Φ_T corresponding to the added elements using the Schur-Banachiewicz inverse formula [9]. Then, the inverse is updated after the estimated signal is pruned by removing the columns that corresponding to the pruned elements. Therefore, large matrix inversion is avoided, resulting in a significant complexity and time reduction. This is especially useful in the case of signals of larger sizes.

Pruning: Next, FMP prunes the estimated signal by removing the elements that have the least contribution to the estimated signal from the identified support set. Only elements which correspond to the k largest magnitude components of the estimated signal are retained while the rest are set to zero, where k is the sparsity of x which is the number of training samples per subject. This operation enhances the reconstruction accuracy and convergence speed of FMP by removing erroneous indices from the identified support set. FMP updates the inverse $(\Phi_i^T\Phi_i)^{-1}$ by removing the pruned columns from Φ_i. This is necessary for the iterative structure of the algorithm, since in the next iteration, we will add new columns. We apply the Schur-Banachiewicz inverse formula [9] but in a reverse manner compared to the estimation phase. More specifically, given the inverse of a matrix $M = \Phi_i^T\Phi_i$, our goal is to obtain the inverse of a matrix $A = \Phi_{i+0.5}^T\Phi_{i+0.5}$ such that $\Phi_i = \begin{pmatrix} \Phi_{i+0.5} & R \end{pmatrix}$, where R is a matrix consisting of the columns to be pruned. Initially, the columns of Φ_i are reordered such that the columns to be pruned are on the right-hand side, while those to be kept are on the left hand side. By defining a matrix N as:

$$N = M^{-1} = (\Phi_i^T\Phi_i)^{-1} = \begin{pmatrix} \Phi_{i+0.5}^T\Phi_{i+0.5} & \Phi_{i+0.5}^T R \\ R^T\Phi_{i+0.5} & R^T R \end{pmatrix}^{-1} \tag{4}$$

Let $A = \Phi_{i+0.5}^T \Phi_{i+0.5}$, $B = \Phi_{i+0.5}^T R$, $C = R^T \Phi_{i+0.5}$, and $D = R^T R$, we can rewrite N as:

$$N = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix} \quad (5)$$

The inverse $(\Phi_i^T \Phi_i)^{-1}$ is computed using the Schur-Banachiewicz formula [9]:

$$(\Phi_i^T \Phi_i)^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}CA^{-1} & S^{-1} \end{pmatrix} \quad (6)$$

where the matrix $S = D - CA^{-1}B$. Thus, we obtain A^{-1} as

$$\begin{aligned} A^{-1} &= (\Phi_{i+0.5}^T \Phi_{i+0.5})^{-1} \\ &= N_{11} - N_{12} N_{22}^{-1} N_{21} \end{aligned} \quad (7)$$

This involves obtaining the inverse of the matrix N_{22}, which is a square matrix of order equal to the number of pruned columns in the i^{th} iteration. However, the dimensions of N_{22} are much smaller than N resulting in a minor inversion overhead due to the significant reduction in computations that is achieved by avoiding the inversion of a much larger matrix.

Since the pruned columns are removed from the inverse $(\Phi_{i+0.5}^T \Phi_{i+0.5})^{-1}$, this inverse can be used in the following iteration. Hence, FMP only performs the matrix inversion $({\Phi_i}^T \Phi_i)^{-1}$ in the first iteration to obtain

$$\Phi_1^\dagger = (\Phi_1^T \Phi_1)^{-1} \Phi_1^T, \quad (8)$$

In successive iterations, $(\Phi_i^T \Phi_i)^{-1}$ is obtained using $(\Phi_{i-1}^T \Phi_{i-1})^{-1}$.

Residual Calculation: Finally, FMP calculates the residual of subtracting the contribution of the pruned estimated signal from the measurement vector. The residual is given by

$$r = y - Z\hat{x} \quad (9)$$

In successive iterations, the residual is correlated with the columns of the sensing matrix. The above FMP steps are repeated until a stopping condition is met. The algorithm terminates if the norm of the residual is less than ϵ_1 or if the difference between the norms of the residuals in two successive iterations is less than ϵ_2, whichever occurs first. Otherwise, the number of allowed iterations is forced to a maximum of k iterations. Figure 2 summarizes the FMP algorithm.

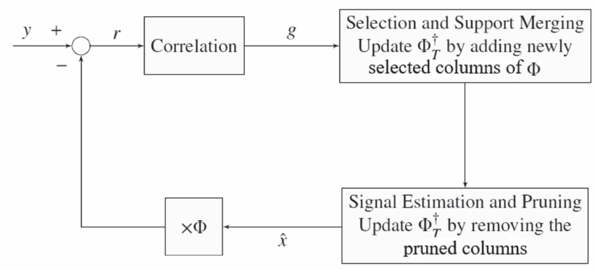

Fig. 2: FMP block diagram.

IV. Simulation Results

For our MATLAB simulations, we use the AR face database [10] and the Extended Yale Face Database B [11]. The AR face database contains more than 4000 colored images. Images are cropped to 120×165 pixels. Images are then converted to grayscale and resized to 128×128 pixels. The Extended Yale Face Database B contains images of 38 persons, 64 images per person, under varying pose and illumination conditions. Images are cropped to 168×192 pixels. We extracted Gabor features from the images and then applied Supervised Locality Preserving Projections (SLPP). All experiments are repeated 10 runs. In each run, one half of the images are randomly selected for training, and the other half for testing [12].

In order to assess the performance gains of using FMP for classification based on sparse representation, we compare it to both ℓ_1 minimization (using homotopy) and OMP [6]. For the FMP algorithm, we set $\alpha = 0.7$ and $\beta = 0.25$ even though the accuracy of the algorithm is not sensitive to the exact values of α and β as we showed in [8]. For the homotopy algorithm, we follow [13].

We plot the recognition rate and time versus the number of used features for the AR database in Fig. 3 and Fig. 4, respectively. As shown in Fig, 3, the three considered reconstruction algorithms result in very high recognition rates that exceed 99%. For instance, the reconstruction rate is about 99.4% at a feature dimension of 100. Furthermore, the recognition rates of FMP and OMP greedy recovery algorithms are very close to that of the optimal and non-practical ℓ_1 minimization. However, using FMP takes the least reconstruction time of about 0.01 seconds at 100 features, compared to 0.02 seconds for OMP, and 0.22 for ℓ_1. Fig. 4 shows that the reconstruction time increases rapidly for ℓ_1 minimization as the dimension increases and reaches about 0.5 seconds for 500 features. While OMP takes 0.08 seconds, our proposed FMP takes only 0.04 seconds at that point. This shows that FMP is capable of accurate reconstruction, considerably faster than other algorithms. More specifically the recognition time of our FMP is only 4% to 8% of that of ℓ_1 depending on the dimension, and approximately 50% of that of OMP.

The Extended Yale Database B shows similar results, illustrated in Fig. 5 and Fig. 6. All algorithms give a reconstruction rate close to 98%. However, FMP achieves a recognition time of about 10% of that of ℓ_1 minimization and about 30% to 50% of that of OMP.

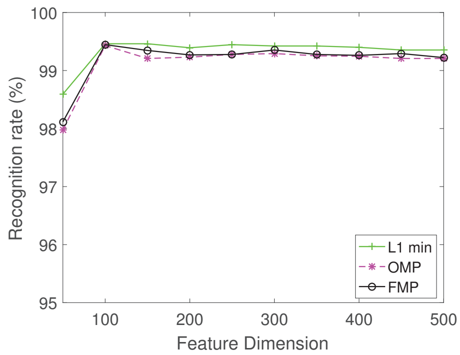

Fig. 3: Recognition rate of the different recovery algorithms (AR database).

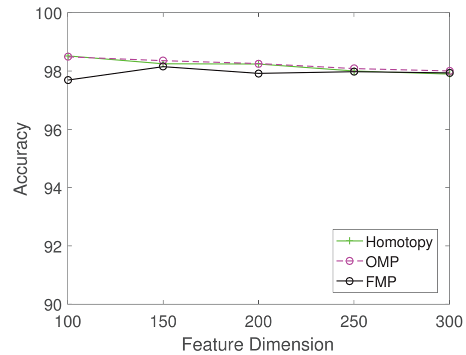

Fig. 5: Recognition rate of the different recovery algorithms (Extended Yale database B).

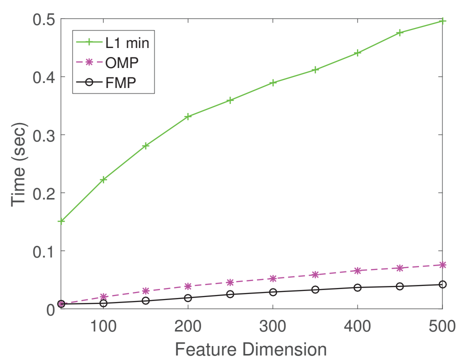

Fig. 4: Reconstruction time of the different recovery algorithms (AR database).

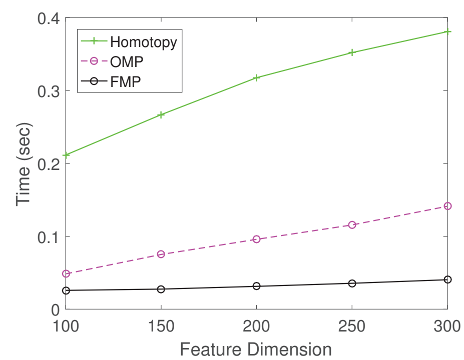

Fig. 6: Reconstruction time of the different recovery algorithms (Extended Yale database B).

V. Conclusion

In this paper, we have presented the Fast Matching Pursuit sparse signal reconstruction algorithm. FMP reconstructs sparse signals fast and accurately. FMP remarkable performance is attributed to two factors: First, FMP element selection strategy only selects an adaptive fraction of the highest magnitude elements from a reduced set of elements. This contrasts to related approaches which perform selection from the whole set. Second, FMP performs matrix inversion in an iterative way that is more efficient and faster than typical matrix inversion. We have applied FMP reconstruction algorithm to sparse representation classification to achieve accurate and fast face recognition. Our simulation results have shown that FMP achieves significant reduction in the reconstruction time with a 98-99% recognition rate.

References

[1] A. K. Jain and S. Z. Li, *Handbook of face recognition*. Springer, 2011.

[2] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

[3] Y. Shen, M. Yang, B. Wei, C. T. Chou, and W. Hu, "Learn to recognise: Exploring priors of sparse face recognition on smartphones," *IEEE Transactions on Mobile Computing*, vol. 16, no. 6, pp. 1705–1717, June 2017.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] D. L. Donoho and Y. Tsaig, "Fast solution of ℓ_1-norm minimization problems when the solution may be sparse," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4789–4812, 2008.

[6] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[7] Š. Vitomir and P. Nikola, "The complete gabor-fisher classifier for robust face recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 1, p. 847680, 2010.

[8] M. M. Abdel-Sayed, A. Khattab, and M. F. Abu-Elyazeed, "Adaptive reduced-set matching pursuit for compressed sensing recovery," in *IEEE International Conference on Image Processing (ICIP)*, 2016.

[9] A. Björck, *Numerical methods for least squares problems*. Siam, 1996.

[10] A. M. Martinez, "The ar face database," *CVC technical report*, vol. 24, 1998.

[11] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.

[12] Z. Zheng, F. Yang, W. Tan, J. Jia, and J. Yang, "Gabor feature-based face recognition using supervised locality preserving projection," *Signal Processing*, vol. 87, no. 10, pp. 2473–2483, 2007.

[13] A. Y. Yang, S. S. Sastry, A. Ganesh, and Y. Ma, "Fast ℓ_1-minimization algorithms and an application in robust face recognition: A review," in *IEEE International Conference on Image Processing (ICIP)*, 2010.