
Image Feature Extraction Using Compressive Sensing

Alaa Eleyan¹, Kivanc Kose², and A. Enis Cetin²

¹ Mevlana University

Electrical and Electronics Engineering Dept., Konya, Turkey

aeleyan@mevlana.edu.tr

² Bilkent University

Electrical and Electronics Engineering Dept., Ankara, Turkey

cetin@bilkent.edu.tr, kkivanc@ee.bilkent.edu.tr

Summary. In this paper a new approach for image feature extraction is presented. We used the Compressive Sensing (CS) concept to generate the measurement matrix. The new measurement matrix is different from the measurement matrices in literature as it was constructed using both zero mean and nonzero mean rows. The image is simply projected into a new space using the measurement matrix to obtain the feature vector. Another proposed measurement matrix is a random matrix constructed from binary entries. Face recognition problem was used as an example for testing the feature extraction capability of the proposed matrices. Experiments were carried out using two well-known face databases, namely, ORL and FERET databases. System performance is very promising and comparable with the classical baseline feature extraction algorithms.

1 Introduction

Reliable automated face recognition is useful in several applications such as security and access control systems. There are many other possible uses for facial recognition that are currently being developed. For example, the technology could be used as a security measure at ATMs and airports in order to intensify security. The same concept could also be applied to computers where facial images would replace passwords in the login process. Given still or video images of a scene, the system should identify or verify one or more persons in the scene using a stored database of faces. The face representation falls into two categories[1]. The first category is global approach or appearance-based, which uses holistic texture features and is applied to the face or specific region of it. Many applied well-known algorithms falls in this category such as principal components analysis (PCA) [2, 3], which is also called eigenfaces [4, 5], linear discriminant analysis (LDA) [6, 7], Gabor wavelet transform[8, 9], and Discrete cosine transform[10]. The second

category is feature-based or component-based, which uses the geometric relationship among the facial features like mouth, nose, and eyes. Wiskott et al. [11] implemented feature-based approach by a geometrical model of a face by 2-D elastic graph. Another example of feature-based was done by independently matching templates of three facial regions (eyes, mouth and nose) and the configuration of the features was unconstrained because the system did not include any geometrical model [12]. In this paper we used the concept of compressive sensing (CS) to generate a random measurement matrix. The CS is based on the fact that we can represent images and signals with a small number of coefficients which, in turn, makes CS powerful as a feature extractor [13] [14]. In our proposed approach, the measurement matrix is different from the random measurement matrices used in most CS problems. We used a matrix containing both zero-mean and nonzero-mean rows. We also compared this matrix with another measurement matrix which is constructed using random binary entries. The measurement matrix will serve as a projection matrix to project image vectors to a new space resulting in feature vector with much shorter length. The matrix with both zero-mean and non-zero mean rows showed superior results using both face databases with various feature vector lengths.

The paper is organized as follows: Section 2 discusses the compressive sensing concept; Section 3 explains the proposed approach. The experimental results and discussions are in Section 4 and then results are concluded at the end of this paper.

2 Compressive Sensing

The Nyquist-Shannon sampling theorem [15] is one of the fundamental theorems in signal processing literature. It specifies the conditions for perfect reconstruction of a continuous signal from its samples. If a signal is sampled with a sampling frequency that is at least two times larger than its bandwidth, it can be perfectly reconstructed from its samples. This approach is very simple to implement however, it is not very efficient in terms of data rates. Sampling the signal according to the Nyquist criteria will end up in large amount of samples, most of which may be thrown away in the later parts of the processing e.g. compression. For example in JPEG compression, first the sampled image is transformed into the DCT domain and then most of the negligible valued (small amplitude) DCT coefficients are thrown away.

Compressed sensing (CS) overcomes this problem by taking compressed measurements [16, 18, 20] from the signal. In a compressive sensing framework, the signal is assumed to be K -Sparse in a transformation domain, such as the wavelet domain or the DCT domain. A signal with length N is K -Sparse if it has at most K non-zero and $(N - K)$ zero coefficients in a transform domain. The case of interest in CS problems is when $K \ll N$, i.e., sparse in the transform domain.

In CS instead of taking individual, regularly spaced samples from the signal, a composition of the values of the signal at some instances is taken. These new samples are called compressed measurements \mathbf{y} , and they are collected as follows

$$\mathbf{y} = \phi \mathbf{x} = \phi \cdot \psi \cdot \mathbf{s} = \theta \cdot \mathbf{s}, \quad (1)$$

where ϕ is the $M \times N$ measurement matrix, $M \ll N$, and \mathbf{s} is the K -sparse transform domain representation of the signal \mathbf{x} in the transform domain represented by ψ . The reconstruction of the original signal \mathbf{x} from its compressed measurements \mathbf{y} cannot be achieved by simple matrix inversion or inverse transformation techniques. A sparse solution can be obtained by solving the following optimization problem:

$$\mathbf{s}_p = \operatorname{argmin} \|\mathbf{s}\|_1 \quad \text{such that} \quad \theta \cdot \mathbf{s} = \mathbf{y} \quad (2)$$

One important characteristic of the measurement matrix ϕ is that it does not need to have a specific structure like transformation matrices or sampling matrices. In fact, in [16, 17, 18], the authors states that the measurement matrix should satisfy the *restricted isometry property (RIP)* for a given number of measurements. They also prove that a random matrix with entries that are i.i.d Gaussian random variables, satisfies the RIP property. Measurement matrix can even be constructed from binary entries [19].

Reconstruction of the original signal from these compressed measurements is another active research field in signal processing and mathematics. Different optimization techniques are frequently used for this purpose. However, for the proposed classification method, we are only interested in the sampling part of the CS framework. Therefore, we will not get into the details of these techniques since the proposed method is related to only the sampling part of the CS framework.

As the perfect reconstruction of the original signal from these compressed measurements is possible, it is also possible to state that, these compressed measurements have descriptive information about the original signal. Therefore, they can used as features in a classification process. In the proposed framework, we are taking compressed measurements from face images using gaussian and binary random measurement matrices and use the measurements as features in the classification. The details of the algorithm is presented in Section 3.

3 Proposed Approach

An illustration of the proposed approach is shown in Fig. 1. The face database is divided into two sets; training set and testing set. Each image in both sets are projected into new space using one of the proposed measurement matrices. After generating the feature vectors of both training set and testing sets, an appropriate classifier is used for classifying each test image to its

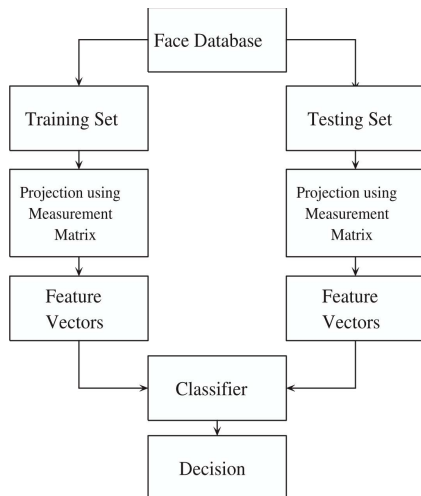


Fig. 1. Flowchart of the proposed approach

corresponding class by comparing its feature vector with the feature vectors of the training set.

The similarity measures used in our experiments to evaluate the efficiency of different representation and recognition methods include ℓ_1 distance measure, δ_{ℓ_1} , ℓ_2 distance measure, δ_{ℓ_2} , and cosine similarity measure, δ_{cos} . The measures for n dimensional vectors are defined as follows

$$\delta_{\ell_1}(x, y) = |x - y| \quad (3)$$

$$\delta_{\ell_2}(x, y) = \|x - y\|^2 \quad (4)$$

$$\delta_{cos}(x, y) = \frac{xy}{\|x\|\|y\|} \quad (5)$$

Experiments were conducted on two commonly used face databases: FERET database [21] and ORL database [22]. For FERET database, 600 frontal face images from 200 subjects are selected. The 600 face images were acquired under varying illumination conditions and facial expressions. Each subject has three images of size 256×384 with 256 gray levels. Each face image is resized to 128×128 . Fig. 2(a) shows sample images from the FERET database. The first two rows are the training images while the third row shows the test images. It can be noticed from this figure that the test images all display variations in illumination and facial expression. To test the algorithms, two images of each subject are randomly chosen for training, while remaining one is used for testing.

The ORL database consists of 400 face images acquired from 40 subjects (i.e., ten images per subject) with variations in facial expression and facial



Fig. 2. Example images from the face databases: (a) Example images from the FERET database. (b) Example images from the ORL database.

details. All images are grey scale with a 92×112 pixels resolution. All images in the database are resized to 128×128 pixels. Fig. 2(b) shows sample images from the ORL database.

4 Comparative Results and Discussions

Preliminary experiments were conducted on both FERET and ORL databases to study the performance of the proposed algorithm for the face recognition problem. Leave-one-out strategy is used in the preparation of the results in Table 1 and 2. Taking p as the number of poses for each person in the database, $p-1$ poses will be used for training while remaining one pose for testing. In this strategy, test pose is changed at every run of the program and the rest $p-1$ is used for training making a total of $(\frac{p!}{p-1!} = p)$ runs. At the end, averaging the results from all these runs will give the final recognition rate.

Both in Table 1 and 2 results of using 3 different measurement matrices were recorded; measurement matrix with zero mean rows, measurement matrix with nonzero mean rows and measurement matrix with mixed zero mean and nonzero mean rows. The measurement matrix size is $M \times N$, where $N = 128 \times 128$ and M can take an arbitrary value which will later represent the resulting feature vector length.

In Table 1, the best performance obtained by using measurement matrix on FERET database with zero mean rows was 81.5%, while it reached 80% by using measurement matrix with nonzero mean rows. Using measurement matrix with zero and nonzero mean rows gave a better performance reached 84.5%. In Table 2, the max performance obtained by using measurement matrix on ORL database with zero mean rows was 96.25%, while it reached 96.5% by using measurement matrix with nonzero mean rows. Using measurement matrix with zero and nonzero mean rows gave a better performance reached 96.75%. It is clear from these results that using a measurement

matrix with both zero and nonzero mean rows helped to slightly improve the performance for both ORL and FERET databases.

Table 3 has the same scenario as in Table 1 and 2. The difference is in the used random measurement matrix. In Table 3 a random measurement matrix with binary entries is used. Normalization of the feature vectors obtained by measurement matrix before the classification stage dropped the performance in Table 1 and 2 by nearly 1 to 2%. On the other hand, normalizing the feature vectors obtained using the binary random measurement matrix helped to improve the performance of the ℓ_1 and ℓ_2 classifiers drastically while it had almost no effect on Cosine distance results. For example, with ORL database, using normalized feature vector of length 1000 gives 96.29% using ℓ_1 distance(Table 3) while it gives 47.5% without feature vector normalization. So, Table 1 and 2 were prepared without normalization of the feature vectors, while Table 3 was prepared using normalized feature vectors.

Even though the results for FERET database show that by using a measurement matrix with both zero and nonzero mean rows (reached 84.5%) can give better performance than using a matrix with binary entries(reached 81%). This improvement was not very clear in ORL database results, as it

Table 1. Face recognition rates on FERET database obtained using different number of features taken using different Measurement Matrices. The rates are given in 3 different metrics defined in (3)-(5).

M	zero mean rows			nonzero mean rows			mixed mean rows		
	δ_{ℓ_1}	δ_{ℓ_2}	δ_{cos}	δ_{ℓ_1}	δ_{ℓ_2}	δ_{cos}	δ_{ℓ_1}	δ_{ℓ_2}	δ_{cos}
50	70.50	70.50	70.00	70.50	70.00	70.00	72.00	71.75	71.50
100	73.00	75.00	72.00	75.00	74.50	75.00	76.00	78.50	75.00
200	78.00	79.50	78.50	79.00	80.50	77.50	80.00	82.00	79.75
300	79.50	79.50	80.00	79.50	80.00	80.00	82.50	82.00	81.50
500	79.00	79.50	79.00	79.75	80.50	81.00	84.50	84.00	81.50
1000	80.50	81.00	81.50	81.00	81.00	80.00	83.00	83.50	82.50

Table 2. Face recognition rates on ORL database obtained using different number of features taken using different Measurement Matrices. The rates are given in 3 different metrics defined in (3)-(5).

M	zero mean rows			nonzero mean rows			mixed mean rows		
	δ_{ℓ_1}	δ_{ℓ_2}	δ_{cos}	δ_{ℓ_1}	δ_{ℓ_2}	δ_{cos}	δ_{ℓ_1}	δ_{ℓ_2}	δ_{cos}
50	89.25	90.25	90.25	91.50	91.50	90.75	90.00	92.75	92.00
100	93.75	93.75	94.00	92.75	93.50	92.75	93.75	94.00	93.75
200	94.50	94.75	93.50	94.75	94.75	94.00	95.00	94.25	94.50
300	95.75	95.00	95.25	95.25	95.00	95.00	95.25	95.75	95.75
500	95.25	96.00	95.50	94.75	95.50	93.50	96.25	96.25	95.00
1000	96.00	96.25	94.50	96.50	95.75	94.50	96.75	96.50	94.75

Table 3. Face recognition rates on FERET and ORL database obtained using different number of features taken using Measurement Matrices with binary entries. The rates are given in 3 different metrics defined in (3)-(5).

M	FERET			ORL		
	δ_{ℓ_1}	δ_{ℓ_2}	δ_{cos}	δ_{ℓ_1}	δ_{ℓ_2}	δ_{cos}
50	69.50	68.25	69.75	90.50	91.00	90.25
100	77.50	76.50	76.75	94.00	95.00	93.00
200	80.50	80.25	80.51	94.75	95.75	94.25
300	80.00	79.75	80.25	94.50	94.25	95.00
500	79.50	78.75	79.25	95.50	95.75	94.25
1000	80.75	80.00	81.00	95.50	96.50	95.25

reached 96.75% by using a measurement matrix with both zero and nonzero mean rows and 96.5% by using a matrix with binary entries, which is a very close performance.

5 Conclusion

In this paper the compressive sensing concept is used to prepare a Gaussian or binary random measurement matrix. Measurement matrix is used as a projection matrix for the image feature extraction. The proposed approach were tested on the face recognition problem. It is experimentally observed that measurement matrices with nonzero mean rows improve results compared to ordinary measurement matrices. This is due to the fact that multiplying an image with a zero mean row is somewhat equivalent to bandpass or highpass filtering. By including nonzero mean rows we also introduce lowpass energy to the measurement process. The preliminary results of the experiments conducted on both FERET and ORL databases indicate that the proposed approach is able to extract the salient features from the face images effectively and provides a high recognition performance . In our future work, more extensive experiments will be carried out on various pattern classification problems to evaluate the performance of the proposed approach under different conditions.

References

1. Chellappa, R., Wilson, C.L., Sirohey, S.: Human and machine recognition of faces: a survey. Proceedings of the IEEE 83(5), 705–741 (1995)
2. Sirovich, L., Kirby, M.: Low-dimensional procedure for the characterization of human faces. Journal of the Optical Society of America A 4(3), 519–524 (1987)
3. Kirby, M., Sirovich, L.: Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Transactions on Pattern Analysis and Machine Intelligence 12(1), 103–108 (1990)

4. Turk, M., Pentland, A.: Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3, 0898-929X, 71–86 (1991)
5. Pentland, A., Moghaddam, B., Starner, T.: Viewbased and modular eigenspaces for face recognition. In: *Proceedings of Computer Vision and Pattern Recognition*, pp. 84–91 (1994)
6. Bellhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(7), 711–720 (1997)
7. Zhao, W., Chellappa, R., Nandhakumarm, N.: Empirical performance analysis of linear discriminant classifiers. In: *Proceedings of Computer Vision and Pattern Recognition*, pp. 164–169 (1998)
8. Shen, L., Bai, L.: A review on Gabor wavelets for face recognition. *Pattern Analysis and Applications* 9(2), 273–292 (2006)
9. Eleyan, A., Ozkaramanli, H., Demirel, H.: Complex wavelet transform-based face recognition. *EURASIP Journal on Advances in Signal Processing*, Article ID 185281, 13 pages (2008)
10. Hafed, Z.M., Levine, M.D.: Face recognition using the discrete cosine transform. *International Journal of Computer Vision* 43(3), 167–188 (2001)
11. Wiskott, L., Fellous, J., Kruger, N., Malsburg, V.: Face recognition by elastic brunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(7), 775–779 (1997)
12. Brunelli, R., Poggio, T.: Face recognition: features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(10), 1042–1052 (1993)
13. Eleyan, A., Kose, K., Cetin, A.E.: New face representation using compressive sensing. In: *IEEE Conference on Signal Processing and Communications Applications*, pp. 558–561 (2011)
14. Liu, L., Fieguth, P., Kuang, G.: Compressed sensing for robust texture classification. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) *ACCV 2010, Part I. LNCS*, vol. 6492, pp. 383–396. Springer, Heidelberg (2011)
15. Shannon, C.E.: Communication in the presence of noise. In: *Proceedings of Institute of Radio Engineers*, vol. 37(1), pp. 10–21 (1949)
16. Baraniuk, R.G.: Compressed sensing (Lecture Notes). *IEEE Signal Processing Magazine* 24(4), 118–124 (2007)
17. Candes, E., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* 52(2), 489–509 (2006)
18. Candes, E.: Compressive sampling. *International Congress of Mathematics* 3, 1433–1452 (2006)
19. Candes, E., Tao, T.: Near optimal signal recovery from random projections: Universal encoding strategies. *IEEE Transactions on Information Theory* 52(12), 5406–5425 (2006)
20. Donoho, D.: Compressed sensing. *IEEE Transactions on Information Theory* 52(4), 1289–1306 (2006)
21. Philipps, P.J., Moon, H., Rivzi, S., Ross, P.: The Feret evaluation methodology for face-recognition algorithms. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 22(10), 1090–1100 (2000)
22. Samaria, F., Harter, A.: Parameterization of a stochastic model for human face identification. In: *Proceedings of the 2nd IEEE Workshop on Applications of Computer Vision*, pp. 138–142 (1994)