MULTI-MODAL BIOMETRICS INFORMATION FROM NUMEROUS UNI-MODAL BIOMETRIC SOURCES

Mhatre Hitesh Kailash Research Scholar, Department of Computer Science Engineering, Kalinga University, Raipur, Chhattisgarh 492101.

Dr. Sanjay Kumar Research Guide, Department of Computer Science Engineering, Kalinga University, Raipur, Chhattisgarh 492101.

ABSTRACT

Biometrics research can be broadly classified into two categories: uni-modal and multi-modal. Multi-modal biometrics is combining information from numerous uni-modal biometric sources. Researchers have indicated that combining information can be beneficial when the quality or information substance of one of the information sources is not adequate for recognition. Various biometric information sources can be combined at different levels; namely, (a) sensor-level, (b) feature-level, score-level, (d) rank-level, and (e) decision-level. Fusion at each level has its advantages and limitations. For example, fusion at the sensor-level can protect most of the information from each of the modalities in any case; sensor-level information may not be extremely discriminatory in nature.

Keywords: Biometrics, multi-model, information, data, recognition, etc.

1. INTRODUCTION

Fusion at each level has its advantages and limitations. For example, fusion at the sensor-level can protect most of the information from each of the modalities in any case; sensor-level information may not be extremely discriminatory in nature. While feature-level fusion doesn't experience the ill effects of noise to the same degree as in the case of sensor-level and also saves significantly more information as compared to score-level, there exist various challenges in utilizing it. To begin with, the relationships between different features are not always known. Second, a few features are variable-length whereas others are fixed-length and therefore concatenation, which is a popular method of feature fusion, is not applicable in countless cases. Third, if these features don't live in a commensurate space it is hard for a classifier to determine reliable decision boundaries. Therefore, relatively less research has zeroed in on feature-level fusion. Multi-modal biometrics can also be beneficial when the data is captured in an unconstrained domain and there are instances of missing information. While researchers have proposed several feature fusion algorithms, not all the algorithms can efficiently combine features within the sight of missing information. The performance of popularly utilized feature fusion algorithms, for example, con-catenation and PCA is significantly affected because of missing information.

2. PRELIMINARIES

In this subsection, we briefly discuss the basic concepts of sparse representation and some recent extension of sparse representation for joint representation and non-linear representation.

2.1 Sparse Representation based Classification

Sparse representation based classification assumes that the training samples of a particular class approximately form a linear basis for a new test sample belonging to the same class. Let vtest be the test sample belonging to the kth class, it can be represented as,

$$v_{test} = \alpha_{k,1}v_{k,1} + \alpha_{k,2}v_{k,2} + \cdots + \alpha_{k,n}v_{k,n} + \epsilon \qquad (1)$$

Where, vk,i denotes the ith training sample and s is the approximation error.

In a classification problem, the training samples and their class labels are provided. The task is to assign the given test sample with the correct class label. This requires finding the coefficients αk,i in Equation 4.1. Since the correct class is not known, SRC represents the test sample as a linear combination of all training samples from all classes,

$$v_{test} = V\alpha + \epsilon \qquad (2)$$

where, $V = \left[\underbrace{v_{1,1}|\ldots|v_{1,n}}_{v_1}|\underbrace{v_{2,1}|\ldots|v_{2,n}}_{v_2}|\ldots\underbrace{v_{c,1}|\ldots|v_{c,n}}_{v_c}\right]$ and

$$\alpha = \left[\underbrace{\alpha_{1,1},\ldots,\alpha_{1,n}}_{\alpha_1},\underbrace{\alpha_{2,1},\ldots,\alpha_{2,n}}_{\alpha_2},\ldots\underbrace{\alpha_{c,1},\ldots|\alpha_{c,n}}_{\alpha_c}\right]$$

According to SRC, only the training samples from the correct class should form the basis for representing the test sample and the samples from other classes should not contribute. Based on this assumption, it is likely that vector α is sparse, i.e., it should have non-zero values corresponding to the correct class and zero values for other classes. Thus Equation 4.2 is a linear inverse problem with a sparse solution. In, the coefficient α is solved by employing a sparsity promoting l1-norm minimization

$$\min_{\alpha} ||v_{test} - V\alpha||_2^2 + \lambda||\alpha||_1 \qquad (3)$$

.

$||\alpha||_1$ denotes the l1 norm of α which $\sum_i^N |\alpha_i|$, where $|\cdot|$ denotes the absolute value function and N represents the length of vector α. *i*

2.2 Block/Joint Sparse Classification

The SRC utilizes a l1-minimization for solving the inverse issue. This is an unsupervised approach and it doesn't use information about the class labels. In it is argued that α should be non-zero for all training samples corresponding to the right class. The SRC assumes that the training samples for the right class will be automatically chosen by imposing the sparsity inducing l1-standard; it doesn't expressly force the constraint that on the off chance that one class is chosen, all the training samples corresponding to that class ought to have corresponding non-zero values in α. claim that it tends to be better recouped if the choice of all the training samples within the class is enforced. This is achieved by employing a supervised l2,1-standard instead of the l1-standard.

$$\min_{\alpha} ||v_{test} - V\alpha|| + \lambda||\alpha||_{2,1} \quad (4)$$

Here, the mixed norm is defined as :

$$||\alpha||_{2,1} = \sum_{k=1}^{n} ||\alpha_k||_2 \quad (5)$$

The inner l2-standard enforces the choice of all the training samples within the class, however the whole of l2-standard over the class' acts as l1-standard over the determination of classes and chooses not many classes. The block sparsity promoting l2, 1-standard guarantees that if a class is chosen, all the training samples within the class are utilized to speak to the test sample.

The Block Sparse representation-based Classification (BSC) approach is successful for general reason classification issues and is appeared to perform well for basic classification issues. Be that as it may, it yields extremely low accuracies compared to SRC for face recognition. To analyze this wonder we allude to Figure 3-2, in BSC all the training samples from the same class have the same class label. Therefore, the l2,1-minimization attempts to choose all the training samples to speak to the test sample. It considers all the hued blocks in Figure 4-2 as a single subspace instead of an association of subspaces; which may not be right approach in all the situations. Enforcing block sparsity is a smart thought when the classification issue is straightforward and all the samples really have a place with a single subspace, for example in fingerprint recognition or character recognition. It keeps choice of samples from arbitrary classes. Be that as it may, face recognition doesn't satisfy this simplistic assumption. As referenced before, the face images can have a place with three sub-spaces. The BSC attempts to combine all the sub-spaces into a single one, for instance, if the test sample is a left profile, it will attempt to fit the left and right profiles as well as the frontal view to the test sample. This is clearly a mistake inclined technique and it has been seen that BSC fails for face recognition related issues, especially in challenging situations with a large variability in the training and test samples.

2.3 Non-Linear Extensions

In non-linear extensions to the SRC and BSC are proposed. The linearity assumption is generalized to include non-linear (polynomial) combinations. The generalization of Equation 4.2 leads to :

$$v_{test} = f(V\alpha) + \epsilon \quad (6)$$

Here, f denotes a non-linear function and s denotes the approximation error. The assumption is that the test sample can be represented as a non-linear combination of the training samples. Notice that this is different from the kernel-based techniques. In these studies, the recovery of the coefficient vector requires solving a non-linear inverse problem with sparsity constraints ,

$$\min_{\alpha} ||v_{test} - f(V\alpha)||_2^2 + \lambda||\alpha||_1 \qquad (7)$$

There are no off-the-rack solutions to comprehend Equation 4.7. In, FOCally Underdeter mined Framework Solver (FOCUSS) and Orthogonal Matching Pursuit (OMP) based solvers are altered to accommodate the non-linearity. The non-linear expansion shows great results on conventional classification issues. Several researchers proposed the Kernel Sparse Representation based Classification (KSRC) approach. KSRC is a straightforward augmentation of the SRC using the Kernel trick. The assumption here is that the non-linear capacity of the test-sample can be represented as a linear combination of the non-linear elements of the training samples, i.e

$$\phi(v_{test}) = \phi(V)\alpha + \epsilon \qquad (8)$$

Here, φ(·) represents a non-linear function. The simplest way to apply the kernel trick is to pre-multiply by φ(V);.

$$\phi(V)^{\intercal}\phi(v_{test}) = \phi(V)^{\intercal}\phi(V)\alpha + \epsilon \qquad (9)$$

The expression in Equation 4.8 consists of inner items between the training samples and the test sample on the left hand side and inner items between the training samples on the correct hand side. When we have the representation as far as inner items, the kernel-trick can be applied as follows,

$$k(x_i, x_j) = \left\langle \phi(x_i), \phi(x_j) \right\rangle \qquad (10)$$

Here, (....) represents the inner product. Applying the kernel trick allows representing Equation 4.9 in the following form

,

$$v_{test}^k = (V)^k\alpha + \epsilon \qquad (11)$$

Here, the superscript k represents the kernelized version of the test sample and training data. Equation 4.11 can be solved using any standard l1-solver. The elegant formulation of the kernel trick we have discussed here was proposed in. In other studies, the sparsity promoting solver (l1-minimization or OMP) was modified to accommodate the kernel trick.

3. PROPOSED GROUP SPARSE REPRESENTATION BASED CLASSIFICATION

The proposed GSRC algorithm is a conventional classification algorithm that can handle various features and data hotspots for each data point. In this research, we propose the formulation and discuss its application for the issue of multimodal biometrics.

Leave N alone the quantity of biometric modalities; for each modality, we assume that the sparse representation classification model remains constant, i.e., the test sample from that modality can be communicated as a linear combination of the training samples of the correct class from the same modality.

$$v^i_{test} = V^i\alpha^i + \epsilon \quad \forall i \in \{1 \dots N\} \qquad (12)$$

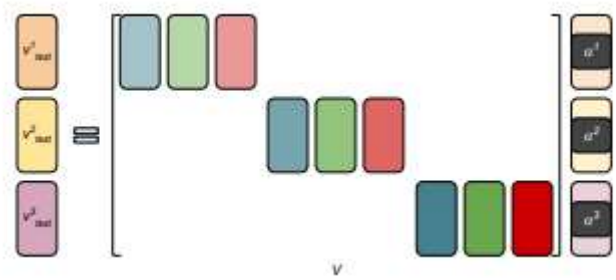


Figure 1: Illustrating the proposed GSRC algorithm.

It is possible to solve each of modalities using the SRC algorithm and combine them at a later stage using a score level fusion rule. However, such an approach does not exploit the intrinsic structure of the problem. A better approach is to combine all the modalities into a single frame- work. As shown in Figure 3-3, this can be succinctly represented as:

$$\begin{bmatrix} v^1_{test} \\ \cdots \\ \cdots \\ v^N_{test} \end{bmatrix} = \begin{bmatrix} V^1 & \dots & 0 \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ 0 & \dots & V^N \end{bmatrix} \begin{bmatrix} \alpha^1 \\ \cdots \\ \cdots \\ \alpha^N \end{bmatrix} + \epsilon \qquad (13)$$

Since each of the α(i)'s are sparse, the simplest way to solve Equation 13 is to impose a sparsity penalty and solve it via l1-minimization. However, such a naive approach is sub-optimal and does not exploit the underlying structure of the problem either

.

$$\alpha = \Big[\underbrace{\alpha^1_1, \dots, \alpha^1_k, \dots, \alpha^1_c}_{\alpha^1}, \dots, \dots, \underbrace{\alpha^N_1, \dots, \alpha^N_k, \dots, \alpha^N_c}_{\alpha^N}\Big]$$

The coefficient vector (represented as a row vector for simplicity) for each modality can be expanded as $\alpha^i = [\alpha^i_1, \dots, \alpha^i_k, \dots, \alpha^i_c]$ where, α_i denotes the coefficients corresponding to the kth class for the ith modality. If a test sample belongs to the kth class, the corresponding coefficients are non-zero. Since the SRC assumptions holds true for individual modalities, the α_i^k 's for each ith (modality) have non-zero values. Therefore, α has a group sparse structure where the non-zero elements occur corresponding to the indices of the kth class. This leads to a group sparse representation where the grouping is simply based on the indices. Equation 3.16 can be solved using the group sparsity promoting l2,1-norm

$$\min_{Z} ||v_{test} - V\alpha||_2^2 + \lambda||\alpha||_{2,1} \quad (14)$$

where,

$$v_{test} = \begin{bmatrix} v_{test}^1 \\ \cdots \\ \cdots \\ v_{test}^N \end{bmatrix}, V = \begin{bmatrix} V^1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ 0 & \cdots & V^N \end{bmatrix} \; and \;\; \alpha = \begin{bmatrix} \alpha^1 \\ \cdots \\ \cdots \\ \alpha^N \end{bmatrix} \quad (15)$$

The proposed GSRC formulation doesn't experience the ill effects of these limitations that fraught block sparse classification. Here, we are not trying to fit one vector (test sample) to all the sub spaces simultaneously (as is finished by BSC); however we are fitting test samples from each modality into the sub spaces spanned by the training samples of the same modality. In less complex words, the main distinction between our suggestion and past examinations is that we define the gathering based on indices from different modalities whereas past investigations define the gathering based on class labels. The l2,1-standard has also been utilized in the sparse representation literature in different ways where its equivalence to hyper-complex sparse coding is leveraged to extract multi-channel quaternionic sparse representations of iris orientation features. The gathering sparsity constraint is applied while optimizing the dictionary coefficient vector for the individual channel encoding. In contrast, in the proposed algorithm, we apply the gathering sparsity constraint on the multi-modal multi-feature representation matrix and enforce bunch sparsity at the index level itself. Our formulation keeps the adaptability of the SRC approach and enhances it by exploiting the multi-modal biometrics issue structure. The representative sample for each class for all the modalities are registered as:

$$v_{rep}(k) = \begin{bmatrix} V_k^i \alpha_k^i \\ \cdots \\ \cdots \\ V_k^N \alpha_k^N \end{bmatrix} \quad (16)$$

The classification is based on the same principle as SRC. The test sample is assigned to the class having the minimum residual error between the test vector and the class representative. One can also use the sum of l1-norm of the αi 's for each class and assign the test sample to the class having the maximum value that is in agreement with the proposal in. Regardless of the criterion (minimum residual or maximum coefficient), GSRC utilizes an elegant decision rule that does not require score level fusion strategies.

k

4. EXPERIMENTAL RESULTS AND ANALYSIS

This section details the databases and experimental protocol, followed by experimental results and analysis. The proposed Group Sparse Representation Classifier or GSRC algorithm is evaluated on two publicly available multimodal biometric databases:

- **WVU multimodal database:** The WVU multi-modal database consists of data pertaining to iris, fingerprint, palm print, hand calculation, face video and voice, and face modalities for 270 individuals. The database also includes soft biometric information, for example, tallness, weight, identity, and sexual orientation. In this research, we center around three biometric modalities: iris, face, and fingerprint. For certain individuals, not all biometric samples are available, these are treated as cases of missing data and information about the concerned (missing) modality is not utilized for recognizing these test images. Images pertaining to 108 subjects (40%) are utilized for training and data for the remaining 162 subjects (60%) is utilized for testing. Three images are utilized as gallery and the remaining images are utilized as tests. The quantity of images available per modality varies and therefore the quantity of test images varies in the range of 680 to 6300.

- **LEA multimodal database:** T he LEA database contains unconstrained multi-modal bio-metric data pertaining to 18,000 individuals. The database comprises of the face, finger-print, and iris modalities. Similar to the WVU database, data for all three modalities is not available for each individual and thus the database encompasses all bio-metric covariates as well as the missing data issue. Data pertaining to half of the individuals, i.e., 9000, is utilized for training and the remaining 9000 individuals are utilized for testing. Two images from each individual are utilized as gallery and the remaining images (1-3 images for every individual) are utilized as tests.

4.1 Algorithms used for Performance Evaluation

In order to evaluate the performance in a multi-feature multimodality setting, two features are considered for each modality. Uniform Circular LBP (UCLBP) and Speeded Up Robust Features (SURF) are considered for face, Video-based Automatic System for Iris Recognition

Table 1 Rank-1 identification accuracy (%) with individual features and their combination (SRC and GSRC) on the WVU and LEA databases

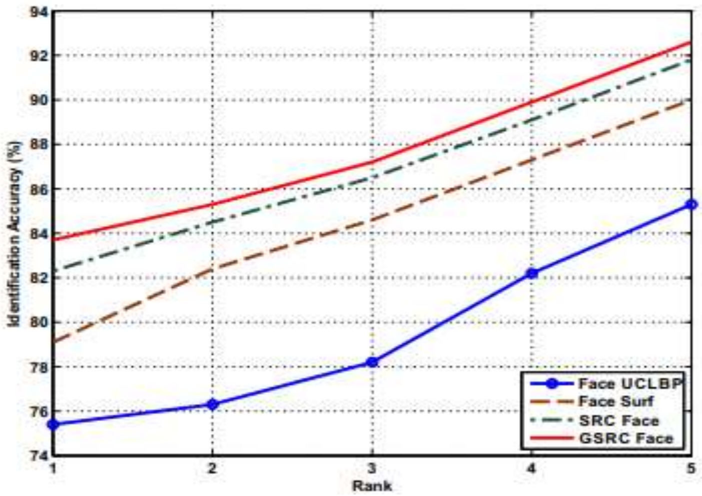
Modality	Features		WVU	LEA
Face	Individual	UCLBP	75.4	24.2
		SURF	79.1	28.4
	Fusion	SRC	82.3	39.7
		GSRC	83.7	40.9
Iris	Individual	Vasir	85.0	31.0
		LPG	90.5	36.4
	Fusion	SRC	92.9	41.2

		GSRC	93.5	43.5
Finger	Individual	NBIS	85.9	40.1
		VeriFinger	90.7	45.7
	Fusion	SRC	92.6	51.8
		GSRC	93.1	53.5

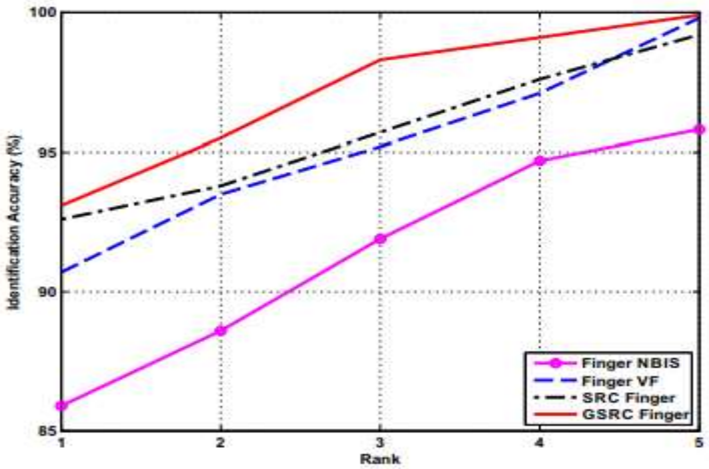
(VASIR) and Log Polar Gabor (LPG) [205] are considered for iris, and NIST Biometric Image Software (NBIS) 1 and VeriFinger (VF)2 are utilized for the fingerprint modality. We utilize the two-stage iris segmentation algorithm proposed in, in which first the inner and external boundaries of the iris are estimated using an elliptical model. Then, the altered MumfordâA˘ S¸ Shah functional is applied in a narrow band over the boundaries estimated in stage one to perform exact segmentation of the iris. The performance of the proposed GSRC algorithm is compared with the SRC algorithm and the state-of-the-art multimodal algorithm which is based on Setting Switching Further, we use total principle match score fusion for performance comparison.

4. 2 Results and Analysis

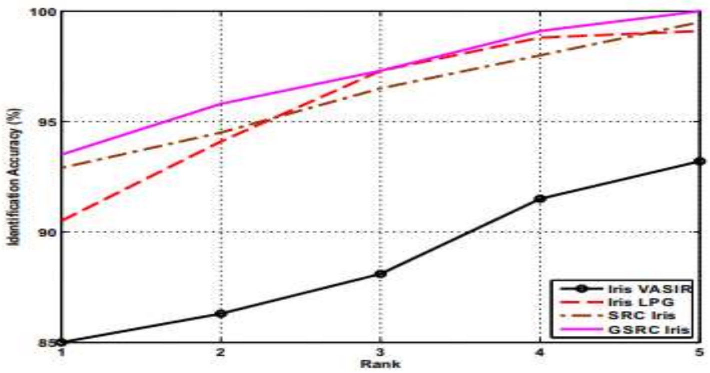
Identification experiments are performed on both the WVU and LEA databases and the performance of the proposed GSRC algorithm is evaluated in four scenarios and major observations are noted below. All the experimental results are presented in the form of CMC bends in Figures summarized in Tables 4.2 and 4.3. Single-feature single-modality: These experiments are performed to assess the baseline performance of the individual features. As referenced before, two features



(a)

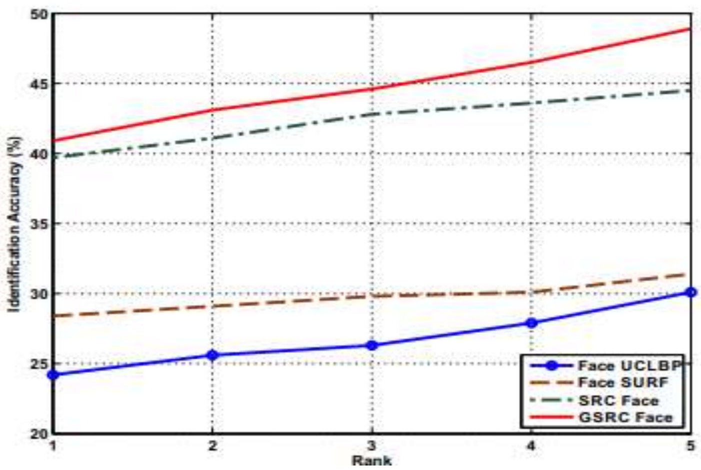


(b)

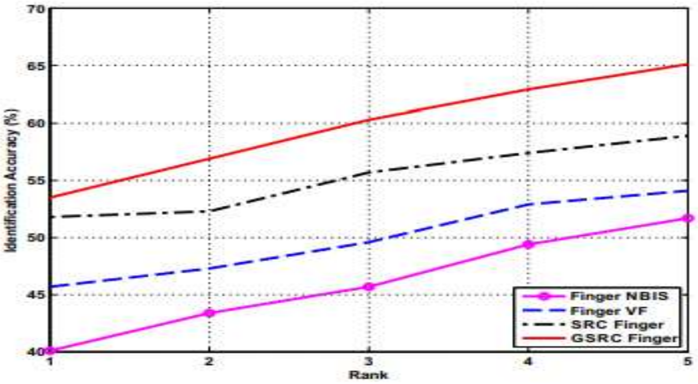


©

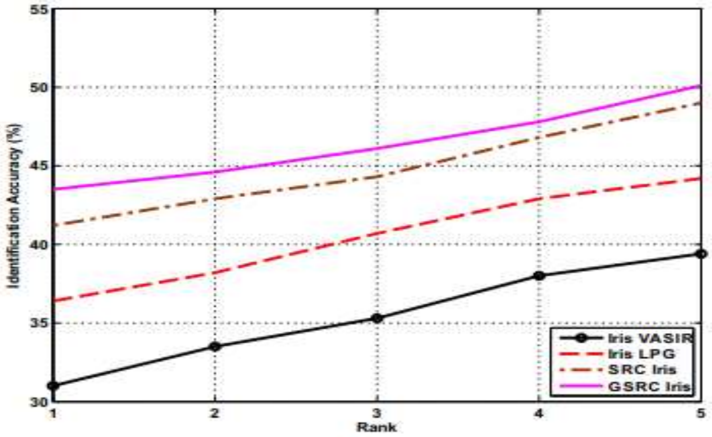
Figure 2: CMC curves on the WVU multimodal database: individual features, SRC and GSRC on (a) face, (b) fingerprint and (c) iris



(a)



(b)



©

Figure 3: CMC curves on the LEA multimodal database: individual features, SRC and GSRC on (a) face, (b) fingerprint, and (c) iris.

Table 2: Rank-1 identification accuracy (%) with fusion of multiple modalities and multiple features (SRC and GSRC) on the WVU and LEA databases

Modality	Fusion Algorithm	WVU	LEA
Face and Iris	SRC	93.9	45.3
	GSRC	94.6	47.4
Face and Finger	SRC	94.4	52.1
	GSRC	95.3	55.8
Iris and Finger	SRC	95.6	52.5
	GSRC	95.9	55.1
	Sum Rule (score level)	95.0	52.6

Face, Finger and Iris	Context Switching [15]	95.8	55.8
	SRC	95.1	54.6
	GSRC	99.1	62.3

For each modality. UCLBP and SURF features are considered for face, VASIR and LPG are considered for iris, and NBIS and VF are utilized for fingerprint. It is seen that the fingerprint and iris modalities perform superior to face on the two databases; nonetheless, fingerprint features outperform the iris modality on the LEA database, potentially denoting the higher reliability of fingerprint modality when data is captured in unconstrained, real-world conditions. It is also seen that no single modality or feature offers the best performance, particularly on the LEA database, clearly motivating the prerequisite for a multi-modal multi-feature recognition algorithm.

5. CONCLUSION

Biometrics is a challenging issue because of various covariates, for example, posture, illumination, and expression, which can adversely influence recognition performance. In any case, utilizing various features to speak to each sample can give strength and enhance the accuracy of recognition algorithms. In this research, we present the gathering sparse representation based classifier for multimodal multi-feature biometric recognition. The proposed algorithm operates on the feature vectors obtained from different modalities/descriptors and perform recognition via feature level fusion and classification. Results on two multimodal databases showcase the productivity of the favorable to presented algorithm in comparison to existing state-of-the-art algorithms. The GSRC algorithm is able to encode the complementary information obtained using different modalities and features to per-form accurate identification in unconstrained scenarios.

REFERENCES

1. Muhammad Sharif et al.: ―Face Recognition: A Survey‖, Journal of Engineering Science and Technology Review 10 (2) (2017) 166- 177

2. Azeem, Aisha, et al. "A survey: face recognition techniques under partial occlusion." Int. Arab J. Inf. Technol. 11.1 (2014): 1-10.

3. Sharif M., Mohsin S., Hanan R., Javed M. and Raza M., "3D Face Recognition using Horizontal and vertical Marked Strips", Sindh University Research Journal (SURJ), 43(01-A), (2011)

4. Wang, Jizeng, and Hongmei Yang. "Face detection based on template matching and 2DPCA algorithm." Imageand Signal Processing, 2008. CISP'08. Congress on. Vol. 4. IEEE, 2008.

5. Zhang, Xiaozheng, and YongshengGao. "Facerecognition across pose: A review." Pattern Recognition42.11 (2009): 2876-2896.

6. Riddhi Patel#1, Shruti B. Yagnik*2 A Literature Survey on Face Recognition Techniques International Journal of Computer Trends and Technology (IJCTT) – volume 5 number 4 –Nov 2013 ISSN: 2231-2803 http://www.ijcttjournal.org Page189

7. S. Tolba, A.H. El-Baz, and A.A. El-Harby, ” Face Recognition: A Literature Review”, International Journal of Signal Processing 2:2 2006

8. Sushma Jaiswal, Dr. (Smt.) Sarita Singh Bhadauria, Dr. Rakesh Singh Jadon," COMPARISON BETWEEN FACE RECOGNITION ALGORITHM-EIGENFACES, FISHERFACES AND ELASTIC BUNCH GRAPH MATCHING", Volume 2, No. 7, July 2011 Journal of Global Research in Computer Science

9. Ming-Hsuan Yang, David J. Kriegman and NarendraAhuja, "Detecting Faces in Images: A Survey," IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 24, NO. 1, JANUARY 2004

10. Li X. and Areibi S.,—"A Hardware/Software codesign approach for Face Recognition", The 16th,International Conference on Microelectronics, Tunisia,2004

11. Yue Ming, Qiuqi Ruan, Xiaoli Li, Meiru Mu, " Efficient Kernel Discriminate Spectral Regression for 3D Face Recognition", Proceedings Of ICSP 2010

12. A.S. Tolba, A.H. El-Baz, and A.A. El-Harby, "Face Recognition: A Literature Review", International Journal of Signal Processing 2; 2 www.waset.org spring 2006

13. Andrea F. Abate, Michele Nappi, Daniel Riccio, Gabriele Sabatino, "2D and 3D face recognition: A survey", PG.NO. 1885–1906, 2007.

14. A.Saradha and S. Annadurai, "A Hybrid Feature Extraction Approach for Face Recognition Systems", ICGST-GVIP Journal, Volume (5), Issue (5), May 2005.

15. Chunghoon Kim, Jiyong Oh, and Chong-Ho Choi, "Combined Subspace Method Using Global and Local Features for Face Recognition" Proceedings of International Joint Conference on Neural Networks, Montreal, Canada, July 31 - August 4, 2005, Page(s): 2030- 2035