Deep multimodal biometric recognition using contourlet derivative weighted rank fusion with human face, fingerprint and iris images

K. Gunasekaran, J. Raja & R. Pitchai

Published online: 14 Jun 2019.

Submit your article to this journal ↗

Article views: 959

View related articles ↗

View Crossmark data ↗

Citing articles: 2 View citing articles ↗

Taylor & Francis
Taylor & Francis Group

SPECIAL ISSUE: COMPUTATIONAL INTELLIGENCE AND CAPSULE NETWORKS   OPEN ACCESS   Check for updates

Deep multimodal biometric recognition using contourlet derivative weighted rank fusion with human face, fingerprint and iris images

K. Gunasekaran[a], J. Raja[b] and R . Pitchai[c]

[a]Research Scholar, Anna University, Chennai, India; [b]Department of ECE, Adhiparasakthi Engineering College, Melmaruvathur, India; [c]Department of CSE, BV Raju Institute of Technology, Narsapur, India

ABSTRACT

The goal of multimodal biometric recognition system is to make a decision by identifying their physiological behavioural traits. Nevertheless, the decision-making process by biometric recognition system can be extremely complex due to high dimension unimodal features in temporal domain. This paper explains a deep multimodal biometric system for human recognition using three traits, face, fingerprint and iris. With the objective of reducing the feature vector dimension in the temporal domain, first pre-processing is performed using Contourlet Transform Model. Next, Local Derivative Ternary Pattern model is applied to the pre-processed features where the feature discrimination power is improved by obtaining the coefficients that has maximum variation across pre-processed multimodality features, therefore improving recognition accuracy. Weighted Rank Level Fusion is applied to the extracted multimodal features, that efficiently combine the biometric matching scores from several modalities (i.e. face, fingerprint and iris). Finally, a deep learning framework is presented for improving the recognition rate of the multimodal biometric system in temporal domain. The results of the proposed multimodal biometric recognition framework were compared with other multimodal methods. Out of these comparisons, the multimodal face, fingerprint and iris fusion offers significant improvements in the recognition rate of the suggested multimodal biometric system.

1. Introduction

The identifiers related to biometric are distinguishing and quantifiable characteristics used to label and trace individual human beings. Certain well-established biometrics used for human identification is face, fingerprint, palm, ear, voice and so on. Many of the real-world biometric systems also referred to as the unimodal; heavily depend on the single source of biometric information. On the other hand, multimodal biometric system observes two or more features from human biometric sample to determine a person's authentication. Multimodal biometric system highly increases the recognition performance. Because, combining multiple pieces of evidence (i.e. multimodal features) for human identification is more effective and reliable. Though, the problem of information fusion still needs to be improved for optimizing the recognition rate of multimodal biometric system.

A general formulation based on Low-Rank and Joint Sparse Representations (LR-JSR) for Multi-Modal Recognition was presented in [1]. A multi-modal feature-level fusion method was designed by applying an optimization algorithm, Alternating Direction Method of Multipliers (ADMM). Next, a modified formulation was structured on the basis of common sparse

and low-rank representation to greatly influence the information related to correlation and coupling beyond the modalities specifically when the performance of each modality differed a lot. Finally, the method was evaluated on several multi-modal recognition problems, resulting in providing the best recognition results. Though the best recognition results were said to be achieved, the performance of each modality is about the same. But, there is no guarantee that enforcing the common representation may always show better results potential solution [fusion – multi modality]

Discriminant Correlation Analysis (DCA) [2] uses class associations of samples in the analysis. The objective of DCA remained in identifying the transformations that increase the pair-wise correlations between two different feature sets. Along with that, DCA also separated the classes within each set. In addition to that, a multiset method was also designed to generalize DCA to be applicable to more than two sets of variables using Multiset Discriminant Correlation Analysis (MDCA).

These characteristics made DCA an effective feature fusion for pattern recognition applications and were also proven to be computationally efficient to be employed in real-time applications. Despite computational efficiency, however, with the varied

---

CONTACT K. Gunasekaran ✉ gunaasekarann@yahoo.co.in 📍 Research Scholar, Anna University, Chennai, India

temporal domain, performance degradation is said to occur by using traditional biometric recognition system. A potential solution is to use a comprehensive deep learning framework.

Multimodal biometric recognition systems present a new way of measuring that in a way provides users' natural behaviour as the core for human–computer interaction. As a result, the research community has initiated to understand how to build a robust and well-integrated multimodal system. But this is said to be achieved only through multimodalities and therefore resulting in a significant commercial impact.

To overcome the shortcoming of software attack [3], another technique based on the fusion of face and iris was presented. However, with varied pose variations, the method was not proved to be efficient. As a result, to address this issue, a novel approach with the aid of hand surface [4] was designed. The novel approach was proved to be efficient even under various poses due to the application of dynamic fusion.

With the increasing need for secure authentication mechanisms, several biometric signals have been analysed in an extensive manner. As a novel approach user authentication that have proved benefits of biometric recognition [5] used faces and gestures obtained from a single vision sensor combined with facial images were adopted. Yet another robust biometric authentication method based on fingerprint was designed using Map and Murillo-Escobar's algorithm [6].

A lot of research work has been done on the multimodal biometric system. A multimodal biometric system using face and fingerprint by introducing Zernike Moment (ZM) and Radial Basis Function (RBF) [7] provided means for better recognition. On the other hand, encryption techniques based on the verification key [8] was designed, therefore improving the storage time and space. Performance evaluation multimodal authentication system based on KNN classifier performed the classification task and operated autonomously [9].

In this work, the multimodal biometrics recognition framework using deep learning techniques with human face, fingerprint and iris images is designed.

The major contribution of the proposed DCD-WR framework is summarized as follows,

(1) Contourlet Transform is applied in DCD-WR framework for performing pre-processing on input face, fingerprint and iris images. Besides, four rank orientations are implemented in contourlet transform in order to obtain most relevant features. This in turns, dimensionality of the feature vector is highly reduced.

(2) Local Derivative Ternary Feature Extraction algorithm is developed in DCD-WR for extracting multimodal features (i.e. face, fingerprint and iris) to reduce computation time by eliminating the redundant information. This helps to improve the recognition accuracy.

(3) A novel weighted rank level fusion where the biometric sample identity is ascertained from the highest ranks returned by the individual biometric. In addition, Deep Learning Template Matching algorithm is designed in DCD-WR to calculate nonlinear activation function and hyperbolic tangent function. The normalized image obtained through sigmoid function and hyperbolic tangent perform a fusion template matching. This in turns, recognition rate is improved multimodal biometric recognition system.

The rest of the paper is organized as follows. Section 2 reviews the related works in the field of multimodal biometric recognition. Section 3 overviews the framework of Deep Contourlet Derivative Weighted Rank (DCD-WR) and goes through details about the framework of biometric recognition framework and performance measures. Section 4 presents our experimental design, data collection, and experimental results. We also evaluate the performance of our system based on the proposed framework. Section 5 concludes our research.

2. Related works

Biometric recognition is considered to be more powerful than the password recognition model as they are invariably analogous with the user. Biometrics also provides certain types of advantages when compared to these security factors, like, non-repudiation, accuracy, and security. Different biometric recognition systems have been investigated on the basis of several biometrics. In the current scenario, many research works disclose that the walking pattern of a person can also be used for the purpose of recognition.

Gabor Wigner transform [10] was used as a feature extraction from which particle swarm optimization was applied to identify the ascendant features, resulting in the minimization of false acceptance rate. Yet another method to reduce the error rate using novel text-based multimodal biometric approach [11] provided an insight into biometric recognition

Machine learning and generalization capability of support vector machine [12] were designed, therefore improving the higher precision of face recognition. A comprehensive multimodal biometric database using mismatched devices, and multiple acquisition environments [13] were analysed. However, with the mobile phone becoming an integral part of the system, a bimodal recognition system [14] was introduced.

A new multimodal biometric identification system based on finger geometry, knuckle print and palm print features of the human hand was presented [15]

using the coarse-to-fine hierarchical method, demonstrating the feasibility and effectiveness. Image quality assessment [16] using iris, fingerprint and face was presented to differentiate between legitimate and impostor samples.

Yet another collaborative face recognition method using social graph model [17] was proposed with the objective of increasing the face recognition rate. However, confidentiality was not said to be improved. To address this issue, cost-sensitive fusion algorithm [18] was presented to obtain maximal generalization performance. However, certain missing modalities pose several threats in the biometric recognition. Neutral Point Substitution (NPS) method [19] was designed to address the missing values while performing multimodal biometric fusion. A unified framework for quality-based fusion of multimodal biometrics was presented in [20].

A two-stage local similarity based classification learning method was developed in [21] to recognize the plant species with higher recognition rate. Local mean clustering was performed to extract the nearest neighbour in each test sample. However, computational complexity was not minimized. In [22], fast periocular authentication was introduced with the aid of feature extraction and matching process for providing fast response in mobile biometric authentication.

A novel plant diseased leaf segmentation and recognition method was designed in [23] to identify the plant leaf diseases in an effective manner. It successfully segments the diseased leaf image in order to extract the features. A coarse-to-fine palmprint recognition method was planned in [24] for extracting invariant features in the palmprint transformation. However, the recognition time was high.

To alleviate the aforementioned problems, this work aims to propose a framework that is able to increase the recognition rate and reduce the computational time and complexity in the presence of high dimension temporal domain by incorporating discriminative features using the multimodal biometric system. The following section explains the proposed framework in detail.

3. Multimodal biometric recognition

In this section, a Deep Contourlet Derivative Weighted Rank (DCD-WR) framework with Face, and Fingerprint and Iris features of human individuals is designed. Figure 1 shows the block diagram of Deep Contourlet Derivative Weighted Rank (DCD-WR) framework to be followed for biometric recognition system.

As shown in Figure 1, Deep Contourlet Derivative Weighted Rank (DCD-WR) framework mainly discusses the fusion of face, fingerprint and iris biometrics using deep weighted rank level fusion. The DCD-WR framework consists of four major blocks namely, pre-processing, feature extraction, fusion, and
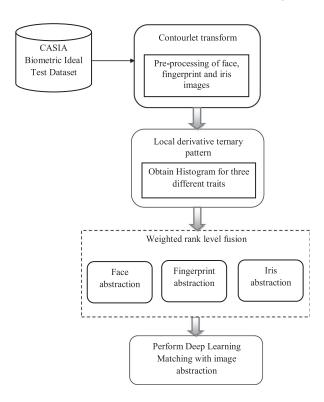

Figure 1. Architecture of deep contourlet derivative weighted rank framework.

classification. The three modalities presumed are face, fingerprint and iris that are given as input. At first, Contourlet Transform is applied to the input face, fingerprint and iris image for performing the pre-processing process. During the pre-processing, the dimensionality reduced features are obtained by using four rank orientation. With the pre-processed output, Local Derivative Ternary Patterns are then applied to obtain a histogram for three different traits. For each pre-processed samples, the feature value (threshold factor) is computed using the local derivative ternary pattern in order to perform efficient recognition. Next, the images are fused using weighted rank level fusion where the weight is assigned represent the significance of each biometric sample. Lastly, deep learning template matching algorithm is used for classification to calculate the recognition accuracy. In deep learning template matching, the sigmoid activation and hyperbolic tangent function are measured to increases the recognition rate.

A Deep Learning Template Matching is developed to find the matching accuracy of test data to the available training data (i.e. from fusion templates) extracted from the benchmark/real dataset. With successful matching of the test and trained dataset (obtained from fusion template), the higher recognition rate is said to be achieved.

3.1. System model for biometric recognition using multimodal

Let us consider a biometric recognition system comprising of "n" users where "$n = 50$ users" and let
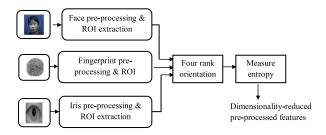
Figure 2. Block diagram of four rank contourlet transform model.

"$\varphi^1, \varphi^2, \ldots, \varphi^n$" denotes the templates of the "*nusers*" in a biometric system. In this setting, "n" biometric features such as face, fingerprint and iris images are registered in different temporal domain. Our problem is then defined as follows: Given a biometric recognition system with an assumption of "$50 users$", the objective lies in designing a framework that aims to improve the recognition rate using multi modal (i.e. face, fingerprint and iris features of human individuals) biometric system.

3.2. Contourlet transform pre-processing

The proposed framework generates a novel feature space in temporal domain using contourlet transform (CNT). The conventional frequency domain transforms are good at separating discontinuities at object edges. However, detecting smoothness along the edges was failed to perform. The contourlet transform effectively overcome the drawbacks of conventional transforms and it obtains smooth contours of images. Because, it performs better while representing the image features like edges, lines, curves. This helps to identify the more related features for image analysis. Therefore, contourlet transforms is better choice especially in the field of image pre-processing.

Contourlet transform is relatively simple. It provides multiresolution images. Contourlet offers a highly efficient image representation on fused image. Contourlet transform obtains the smooth contours of images with the aid of four rank orientations. The entropy based sub-band selection is carried out to choose the most important sub-bands which successfully minimize the dimension of the features. This helps to reduce the computational complexity in contourlet transform.

The proposed framework generates a novel feature space in temporal domain using contourlet transform (CNT). The coefficient of CNT in temporal domain is used as the extracted features. For pre-processing, Four Rank Contourlet Transform is applied over the face, fingerprint and iris images in temporal domain to extract essential features needed for identification. Figure 2 shows the block diagram of Four Rank Contourlet Transform model.

As shown in Figure 2, four rank contourlet transformation is applied to pre-processing the multimodal

features (i.e. face, finger print and iris images are given as input). Contourlet transform uses a four rank orientation in order to get the smooth contours of images. The contourlet transformation reduces the dimension of the feature vectors by means of selecting important features in multimodal biometric recognition. This is achieved with the aid of measuring the entropy of each detail sub-bands (i.e. input). In addition, threshold is determined experimentally to select the most relevant features to recognize the images. This in turns, the computational complexity is considerably reduced in DCD-WR framework. In Four Rank Contourlet Transform, the face, fingerprint and iris images are split into "2^1" orientation at the first rank, "2^2" orientation at the second rank, "2^3" orientation at the third rank and "2^4" correlation at the fourth rank. Hence a total of 30 (i.e. $2^1 + 2^2 + 2^3 + 2^4 = 2 + 4 + 8 + 16 = 30$) orientation sub-bands are obtained for each training image. A threshold is then obtained in an experimental manner to select the prime sub-bands that are employed for feature extraction. This is performed by measuring the Entropy and is mathematically formulated as given below.

$$E = -\sum_{i=1}^{n} w_i \log w_i \qquad (1)$$

From Equation (1), "w_i", symbolizes the histogram weight for "i" temporal domain images, where series of images (i.e. face, fingerprint and iris) are taken at different time. In the proposed framework, with four rank contourlet transform considered, the entropy of 30 different sub-bands at different levels is calculated using the images of CASIA Biometric Ideal Test Dataset. The pseudo code representation of Four Rank Contourlet Transform Pre-processing is given below.

Algorithm 1 Four rank contourlet transform pre-processing.

Input: Biometric sample "i", Fingerprint Impression "$FP_i = fp_1, fp_2, \ldots, fp_n$", face "$F_i = f_1, f_2, \ldots, f_n$", Iris "$R_i = r_1, r_2, \ldots, r_n$",
Output: Pre-processed image "FP'_i, F'_i, R'_i"
1: **Begin**
2:     **For** each biometric sample "i" images with Fingerprint Impression "FP_i", face images "FP_i" and iris images "R_i"
3:        Measure entropy using Equation (1)
4:     **End for**
5: **End**

As given above, using the Four Rank Contourlet Transform Pre-processing, for each biometric sample, multimodal images like, face, fingerprint and iris images are taken as input. To the input images, pre-processing is said to be performed using four rank orientations. By applying four rank orientations, the
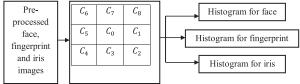
Figure 3. Local ternary pattern for "C_0" with eight neighbours.

threshold factor is measured based on the entropy value with which only the important sub-bands or features for each biometric samples are selected and used for feature extraction.

3.3. Local derivative ternary feature extraction

Followed by the pre-processed features obtained, in this work, features are extracted from the temporal domain based on the Local Derivative Ternary Pattern. The idea has been adopted from the Discriminant Correlation Analysis (DCA) [1] methods where Multiset Discriminant Correlation Analysis (MDCA) was used on two sets of variables at a time. Unlike MDCA adopted in DCA, the Local Derivative Ternary Pattern (LDTP) model does not threshold the pixels into "0" and "1". But, LDTP model uses a threshold constant to threshold pixels into three values "1", "0" and "−1", to which the first order derivative is applied. Figure 3 shows an example of Local Ternary Pattern for "C_0" with eight neighbours "C_1", "C_2", "C_3", "C_4", "C_5", "C_6", "C_7", "C_8", respectively.

As shown in Figure 3, let us consider "Th" as the threshold and "C_0" as the centre pixel for pre-processed image "FP_i', F_i', R_i'", with the value of the neighbouring pixels being "$C_i, i = 1, 2, .., 8$", the result of the threshold factor for LDTP is mathematically formulated as given below.

$$F\begin{pmatrix} i(C_o) \\ i(C_i) \end{pmatrix} = \begin{cases} 1, If i(C_i) > i(C_o) \\ 0, If i(C_i) = i(C_o) \\ -1, If i(C_i) < i(C_o) \end{cases} \quad (2)$$

With the above centre pixel "C_0" with eight neighbouring pixels "C_i", the first order derivative is mathematically formulated as given below.

$$i_0 C_0 = |i(C_0) - i(C_4)| \quad (3)$$

From Equations (2) and (3), by applying Local Derivative Ternary Pattern, reduced coefficients are used as features vectors to recognize the images in a more accurate manner. Therefore, by applying LDTP model, the feature discrimination power is enhanced by selecting the coefficients that has maximum variation across pre-processed multimodality features. The pseudocode representation of Local Derivative Ternary Feature Extraction is given below.

As given in the above Feature Extraction algorithm, the threshold constant to threshold pixels are split into

Algorithm 2 Local derivative ternary feature extraction algorithm.

Input: Pre-processed image "FP_i', F_i', R_i'", Threshold factor "Th",
Output: Optimal features extracted "FP_i'', F_i'', R_i''"
1: **Begin**
2:     **For** each pre-processed image "FP_i', F_i', R_i'" with "i" samples
3:         Obtain result of the threshold factor using Equation (2)
4:         Measure first order derivative using Equation (3)
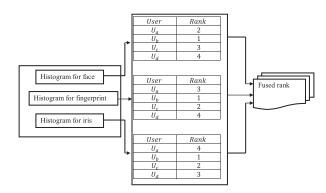5:   **End for**
6: **End**

Figure 4. Example of weighted rank level fusion with four users.

three values using Local Derivative Ternary model. The results of the threshold factor are considered as the feature value for pre-processed multimodal (i.e. face, fingerprint, and iris) samples. The extracted feature value is used to construct a biometric recognition for representing multimodal and performs well under variations of the temporal domain, expression, and inclusion of occlusion.

3.4. Weighted rank level fusion

Local Derivative Ternary Pattern features are extracted for each face, fingerprint and iris image, and the acquired features are fused by using weighted rank level fusion and stored in a database for matching. Figure 4 shows the block diagram of rank level fusion followed in the proposed framework.

Figure 4 shows the example of weighted rank level fusion. As illustrated in the figure, the histogram for face, fingerprint and iris are provided as input for obtaining the fused rank. The interpretation of multimodal biometric samples is not consistent. For example, the biometric sample using face images is anticipated to interpret better than those biometric samples using fingerprint or face images and vice versa. Therefore weighted rank level fusion is used by

allocating corresponding weights to the ranks produced by individual biometric samples has been proposed. The fused rank scores in weighted rank level are computed as follows:

$$FR_k = \sum_{i=1}^{n} w_i r_i(k) \qquad (4)$$

As illustrated in Figure 4, the weighted rank level fusion is employed for integrating the output ranks from the various biometrics users in order to discover the identity of an individual with more accurate. At first, the weighted rank level fusion is employed to produces the matching scores for four different users. Then, the generated matching scores are internally arranged for providing ranking results among the different histogram face, fingerprint and iris identities. There are only four users (User U_a, U_b, U_c, U_d) and 1, 2, 3, 4 denotes the ranks for the input histogram features with 1 being the highest rank. This helps to increase the matching accuracy in DCD-WR framework.

From Equation (4), "w_i" represents the weights assigned to the "ith" multimodal biometric samples. In the weighted rank level fusion, the biometric sample identity is ascertained from the highest ranks returned by the individual biometric samples. The weights are allocated to indicate the consequential of each biometric sample and are evaluated from the overall assessment of the performance. The weights are computed during the training phase by measuring the entropy (1).

3.5. Deep learning template matching algorithm

Finally, with the resultant weighted rank level fusion performed, a Deep Learning Template Matching algorithm is designed to identify the matching accuracy of test data to the available training data. The matching tasks for multimodal biometrics (i.e. face, fingerprint, and images) are carried out using Deep Learning Template Matching algorithm. While considering high definition data, the proposed work is able to provide better efficiency. This is because of DCD-WR framework performs deep learning template matching. Deep learning template matching in biometric multimodal recognition (i.e. face, fingerprint and iris) uses a cascade of several layers of nonlinear processing units for learning and matching the features from the high definition data. Therefore, DCD-WR framework performs better in high definition data.

To perform Deep Learning Template Matching, nonlinear activation function, i.e. sigmoid function and hyperbolic tangent (tanh) function, are measured for each biometric samples. The sigmoid activation function is mathematically represented as given below.

$$S(F_i^{''}) = \left(= \frac{1}{1 + \exp(-w_i T(F_i^{''}) - b_{F_i^{''}})} \right) \qquad (5)$$

$$S(FP_i^{''}) = \left(= \frac{1}{1 + \exp(-w_i T(FP_i^{''}) - b_{FP_i^{''}})} \right) \qquad (6)$$

$$S(R_i^{''}) = \left(= \frac{1}{1 + \exp(-w_i T(R_i^{''}) - b_{R_i^{''}})} \right) \qquad (7)$$

From Equations (5)–(7), the sigmoid activation function for face, fingerprint and iris "$S(F_i)$", "$S(FP_i)$", "$S(R_i)$" is obtained using the input face, fingerprint, iris "F_i", "FP_i", "R_i", weight of face, fingerprint, iris "w_i", and the bias value of face, fingerprint and iris "b_F", "b_{FP}", "b_R", respectively. The hyperbolic tangent function for each biometric sample is mathematically evaluated as given below.

$$T(F_i^{''}) = \exp(W_F T(F_i^{''}) + b_{F_i^{''}}) \qquad (8)$$

$$T(FP_i^{''}) = \exp(W_F T(FP_i^{''}) + b_{FP_i^{''}}) \qquad (9)$$

$$T(R_i^{''}) = \exp(W_F T(R_i^{''}) + b_{R_i^{''}}) \qquad (10)$$

Finally, the multimodal biometric recognition is obtained by measuring the Log Likelihood Ratio (LLR) between the probability "$Prob(U_i, U_j | H_s)$" that the two biometric samples are same or "$Prob(U_i, U_j | H_d)$" that the two biometric samples are different.

$$LLR(U_i, U_j) = \log \frac{Prob(U_i, U_j | H_s)}{Prob(U_i, U_j | H_d)} \qquad (11)$$

In this way, by measuring the LLR values using Equation (11), the higher recognition rate is said to be achieved. The pseudocode representation of Deep Learning Template Matching algorithm is as given below.

Algorithm 3 Deep learning template matching.

Input: Optimal features extracted $FP_i^{''}, F_i^{''}, R_i^{''}$
Output: Improved recognition rate
1: **Begin**
2:     **For** each Optimal features extracted $FP_i^{''}, F_i^{''}, R_i^{''}$
3:         Measure sigmoid activation function for face, fingerprint and iris using Equation (5), (6) and (7)
4:         Measure hyperbolic tangent function for face, fingerprint and iris using Equation (8), (9) and (10)
5:         Measure log likelihood ratio using Equation (11)
6:     **End for**
7: **End**

As shown in the above Deep Learning Template Matching algorithm includes 50 user's unique face,

fingerprint and iris abstraction separately extracted from benchmark/real dataset. Face, fingerprint and iris abstraction includes 50 faces, fingerprints, and iris extracted from 50 individual users' and stored as fusion templates form Deep Learning Template Matching. The multimodal biometric system for the proposed framework is developed by integrating three traits, i.e. face, fingerprint and iris. Based on the sigmoid activation and hyperbolic tangent function, the matching score is evaluated via, LLR. These LLR scores finally improve the recognition rate.

4. Experiments

In this paper, we aim to improve the biometric recognition and reduce the computational time involved. CASIA Biometric Ideal Test Dataset containing 500 face, fingerprint and iris images is used to evaluate the performance of our proposed framework. The face, fingerprint and iris images are selected from CASIA Biometric Ideal Test Dataset, which provides 100 images for each person capturing every combination of features using MATLAB simulator. The biometric samples are extracted from the CASIA Image Database that includes many iris, face, fingerprint, palm print, multi-spectral palm and handwriting image for biometric recognition. In this work, face, fingerprint and iris images are taken as input.

By using CASIA Biometric Ideal Test Dataset and the defined testing framework results are compared with the existing method. Deep Contourlet Derivative Weighted Rank (DCD-WR) is compared with the existing Joint Sparse Representations (LR-JSR) for Multi-Modal Recognition [1] and Discriminant Correlation Analysis (DCA) [2]. The proposed work plan to conduct an experimental and analytical evaluation of multimodality biometric recognition with data sets extracted from research repositories. The experimental evaluation of proposal work is conducted on various factors such as computational time, computational complexity and recognition rate to a different number of human biometric sample.

From Figure 5, the sample 1 has the dimensions of 640*480 and resolutions of 94 dpi and sample 2 have the dimensions of 114*126 and resolutions of 72 dpi. The third sample has the dimensions of 106*114 and resolutions of 96 dpi and fourth sample have the dimensions of 320*240 and resolutions of 72 dpi. Lastly, fifth sample has the dimensions of 768*576 and resolutions of 72 dpi.

The computational time involved during feature extraction is the time required to measure the attributes with respect to the number of human biometric samples. In other words, it is the product of a number of human biometric samples considered and the time taken for feature extraction.

$$CT = n * Time(feature\ extraction) \qquad (12)$$

From Equation (12), the computational time is obtained with respect to the number of human biometric samples "n". Followed by computational time, computational complexity is measured. Computational complexity is calculated while performing the template matching for three different modal, face, iris and fingerprint over different time intervals. It is the amount of working storage required to perform Deep Learning Template Matching algorithm (algorithm). In other words, computational complexity measures the memory required to execute the algorithm at any point.

$$CC = \sum_{i=1}^{n}(n * Mem(LLR(U_i, U_j))) \qquad (13)$$

From Equation (13), the computational complexity "CC" is obtained using the number of human biometric samples "n" and "$Mem(LLR(U_i, U_j))$" refers to the memory consumed while performing log likelihood ratio (face, fingerprint and iris) between two users "U_i, U_j", respectively. It is measured in terms of Kilo Bytes (KB). After that, the recognition rate "R" for biometric recognition using multimodal features depends on the accuracy of number of samples correctly recognized and is evaluated by the following formula.

$$R = \left[\frac{Fusion\ template\ correctly\ matched}{n} * 100\right] \qquad (14)$$

From Equation (14), the matching rate "R" is measured with respect to the total number of human biometric samples "n" in terms of percentage (%). With the measure of recognition rate the recognition accuracy also get improved in DCD-WR framework. Then, true positive rate "TPR" is measured as the ratio of number of biometric sample accurately or correctly recognized to the total number of samples. IT is mathematically calculated as,

$$TPR = \frac{Number\ of\ biometric\ samples\ correctly\ recognized}{No.of\ samples} * 100 \qquad (15)$$

From Equation (15), the true positive rate "TPR" is measured with respect to the different number of samples. It is measured in terms of percentage (%). Lastly, false acceptance rate "(FAR)" in biometric recognition is used to compute the average number of false acceptances to the total number of biometric sample.

$$FAR = \frac{Probability\ of\ incorrect\ recognition\ of\ biometric\ sample}{n} * 100 \qquad (16)$$

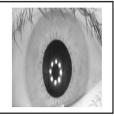From Equation (16), false acceptance rate "FAR" is measured in terms of percentage (%).

Figure 5. Five samples of biometric recognition.

Table 1. Tabulation for computational time.

Number of human biometric sample	Computational time (ms)		
	Proposed DCD-WR	Existing LR-JSR	Existing DCA
50	15.3	19.2	21.5
100	18.2	22.4	26.4
150	21.3	26.3	30.4
200	25.4	27.2	36.8
250	28.3	29.1	39.5
300	31.4	35.2	44.5
350	36.2	40.4	48.2
400	44.1	46.3	52.3
450	48.3	52.6	55.5
500	49.2	55.8	59.3

5. Discussion

In this section, we introduce the performance evaluation based on the implementation by MATLAB simulator. The validation results are presented in three tables. Table 1 represents the computational time for biometric recognition with different number of human biometric samples using Matlab simulator and comparison is made with two other methods, namely LR-JSR [1] and DCA [2].

5.1. Impact of computational time

The first set of experiments is intended to evaluate the computational time performance of the proposed DCD-WR. In the experiment, to perform a precise comparison of the features of both Deep Contourlet Derivative Weighted Rank (DCD-WR) and existing Low-Rank and Joint Sparse Representations (LR-JSR)

for Multi-Modal Recognition [1] and Discriminant Correlation Analysis (DCA) [2], the computational time involved for human recognition is simplified as given below. Followed by sample calculation, Table 1 shows a comparison table of computational time.

Sample calculation

- **LR-JSR:** With "50" as the human biometric sample provided as input and the time for feature extraction being "0.384ms", the computational complexity is obtained as given below.

$$CT = 50 * 0.384ms = 19.2ms$$

- **DCA:** With "50" as the human biometric sample provided as input and the time for feature extraction being "0.43ms", the computational complexity is obtained as given below.

$$CT = 50 * 0.43ms = 21.5ms$$

- **Proposed DCD-WR:** With "50" as the human biometric sample provided as input and the time for feature extraction being "0.306ms", the computational complexity is obtained as given below.

$$CT = 50 * 0.306ms = 15.3ms$$

Table 1 represents the computational time using MATLAB simulator and comparison is made with two other methods, namely LR-JSR [1] and DCA [2]. The information about the number of human biometric samples considered for experimentation is summarized in

Table 1. Higher, the number of human biometric sample, higher the computational time is. This is because with higher biometric sample size, the size of individual images grows exponentially, and therefore the computational time for increased biometric sample size also increases. But from the table, it is evident that the computational time for feature extraction is comparatively observed to be lower using the proposed DCD-WR framework.

By applying Local Derivative Ternary Pattern (LDTP) model in DCD-WR framework, face, fingerprint and iris images are extracted, comparing their histogram threshold value. This, in turn, removes the irrelevant information present in the face, fingerprint and iris images resulting in minimizing the computational time. The process is repeated with human biometric sample size of 50–500 for conducting experiments. The results reported here confirm that with the increase in the number of human biometric samples, the computational time also increases, though betterment achieved using DCD-WR framework.

As shown in the table, when compared to two other methods LR-JSR [1] and DCA [2], the DCD-WR framework had better changes using the extensive Local Derivative Ternary Feature Extraction algorithm. This is because the Local Derivative Ternary Feature Extraction algorithm applied in DCD-WR framework enhances the feature discrimination by selecting the coefficients possessing maximum variation across pre-processed multimodality features reducing certain amount of noise present in face, fingerprint and iris. This, in turn, reduces the computational time by 11% compared to LR-JSR and 25% compared to DCA.

5.2. Impact of computational complexity

Next, we address the second goal of the experiments with respect to computational complexity showing the comparison between DCD-WR framework, LR-JSR [1] and DCA [2]. Computational complexity is computed as the product of number of human biometric sample into memory consumed while performing log likelihood ratio between two users. Ten unique experiments were conducted. Followed by sample calculation, Figure 5 shows the tabulation for computational complexity that performs the template matching for three different modal, face, iris and fingerprint over different time intervals.

Sample calculation

- **LR-JSR:** With the number of human biometric sample being "50", the memory consumed while performing log likelihood ratio between two users is "24.3*KB*", then the computational complexity is obtained as given below.
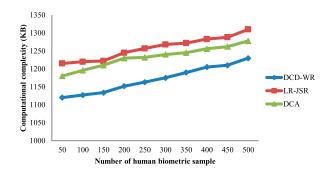
$$CC = 50 * 24.3KB = 1215KB$$

Figure 6. Measure of computational complexity.

- **DCA:** With the number of human biometric sample being "50", the memory consumed while performing log likelihood ratio between two users is "23.6*KB*", then the computational complexity is obtained as given below.

$$CC = 50 * 23.6KB = 1180KB$$

- **Proposed (**DCD-WR**):** With the number of human biometric sample being "50", the memory consumed while performing log likelihood ratio between two users is "22.4*KB*", then the computational complexity is obtained as given below.

$$CC = 50 * 22.4KB = 1120KB$$

The computational complexity is presented in Figure 6 with respect to 500 different human biometric samples. With respect to the increasing number of human biometric samples, the computational complexity is increased, but shows gradual improvement by applying DCD-WR when compared to LR-JSR and DCA. To better perceive the efficacy of the proposed DCD-WR framework, substantial experimental results are conducted and illustrated in Figure 6.

For all scenarios as shown in Figure 6, the computational complexity is increasing with biometric samples considered from different users. The targeting results of computational complexity using DCD-WR framework is compared with two state-of-the-art methods [1], [2] in Figure 6 is presented for visual comparison. Our framework differs from the LR-JSR [1] and DCA [2] in that we have incorporated weighted rank level fusion for multimodalities integrating face, fingerprint and iris of the human individuals.

Here the dominant attributes are stored using Four Rank Contourlet Transform that in turn selects the prime sub-band for feature extraction and therefore reduces the dimensionality factor. With the resultant feature vectors obtained for face, fingerprint and iris image, with high feature discrimination made by applying Local Derivative Ternary Pattern to the resultant normalized image based on the mutual information of the individual features, further reduces the computational complexity during human recognition. Therefore

the computational complexity for biometric recognition using multimodal features (face, fingerprint, and iris) is reduced by 7% compared to LR-JSR [1] and 5% compared to DCA [2], respectively.

5.3. Impact of recognition rate

The main goal of our experiments is to determine the rate of recognition for biometric recognition using machine learning algorithms. 500 face, fingerprint and iris images were randomly selected. Recognition rate is one of the performance metrics to measure the accuracy of the biometric samples that were recognized correctly to the number of human biometric samples provided as input during experimentation

The table 2 shows the results of recognition rate of an exploratory experimentation on CASIA Biometric Ideal Test Dataset by presenting the recognition rate using DCD-WR framework, LR-JSR [1] and DCA [2]. The experiment was conducted to gain insights on the recognition results of the datasets, to measure the performance of recognition rate and to measure the accuracy of biometric recognition using multimodal features. Followed by sample calculation, Table 2 presents the recognition rate obtained for 50 different human biometric samples.

Sample calculation

- **LR-JSR:** Out of 50 human biometric samples given as input, fusion template accurately or correctly matched was 27. These values when substituted in the above formula obtain the recognition rate.

$$R = \frac{27}{50} * 100 = 54\%$$

- **DCA:** Out of 50 human biometric samples given as input, fusion template correctly matched was 33. These values when substituted in the above formula obtain the recognition rate.

$$R = \frac{33}{50} * 100 = 66\%$$

- **Proposed (DCD-WR):** Out of 50 human biometric sample given as input, fusion template correctly matched was 41. These values when substituted in the above formula obtain the recognition rate.

$$R = \frac{41}{50} * 100 = 82\%$$

The comparison of recognition rate is presented in Table 2 with respect to the human biometric samples in the range of 50–500. With the increase in the human biometric samples, initially the recognition rate increases for 200 samples, then a decrease in recognition rate is observed for 250 samples and further improvement is observed from 300 to 500 samples. This

Table 2. Tabulation of recognition rate.

Number of human biometric sample	Recognition rate (%)		
	DCD-WR	LR-JSR	DCA
50	82	54	66
100	84	56	67
150	89	58	68
200	85	57	66
250	86	60	70
300	88	62	73
350	91	64	75
400	92	68	77
450	95	69	78
500	96	70	79

is due to the presence of noise in the face, iris and fingerprint images that in turn affects the recognition rate.

The decision point of 10 different biometric samples (face, fingerprint and iris) was selected in a random manner that achieved a substantial improvement in ratings from the previous decision. The results show better performance of the proposed DCD-WR framework, but however not seen to be linear due to the presence of certain irrelevant information, not removed during the pre-processing stage. The last values of the graph plotted in the figure seem to confirm the working hypothesis that the recognition rate for biometric recognition using multimodalities increases with the increase in the human biometric samples.

As illustrated when compared to two other methods LR-JSR [1] and DCA [2], the DCD-WR framework substantially improved the recognition rate for biometric recognition using the extensive Deep Learning Template Matching algorithm. This is because the DCD-WR framework adapted a sigmoid activation and hyperbolic tangent function separately for each trait (i.e. face, fingerprint, and iris) to decide upon the biometric recognition, resulting in the improvement of the recognition rate. Furthermore based on the Log Likelihood Ratio, matching between the training and test set was conducted based on the mutual information values (i.e. face, fingerprint and iris), that in turn improved the recognition rate by 44% compared to LR-JSR and 24% compared to DCA. As a result, the proposed DCD-WR framework is better to provide more accurate recognition than others methods.

5.4. Impact of true positive rate

Another experiment is focused to evaluate the performance of the true positive rate in proposed DCD-WR. In the experiment, the comparison both Deep Contourlet Derivative Weighted Rank (DCD-WR) and existing Low-Rank and Joint Sparse Representations (LR-JSR) for Multi-Modal Recognition [1] and Discriminant Correlation Analysis (DCA) [2] using true positive rate is given below. Sample calculations are given below.
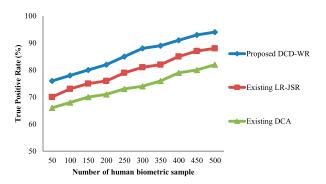
Figure 7. Measure of true positive rate.

Sample calculation

- **LR-JSR:** The 50 human biometric samples given as input, then the biometric samples correctly recognized as 35. Then the true positive rate is calculated as,

$$TPR = \frac{35}{50} * 100 = 70\%$$

- **DCA:** The 50 human biometric samples given as input, then the biometric samples correctly recognized as 33. Then the true positive rate is calculated as,

$$TPR = \frac{33}{50} * 100 = 66\%$$

- **Proposed (DCD-WR):** The 50 human biometric samples given as input, then the biometric samples correctly recognized as 38. Then the true positive rate is calculated as,

$$TPR = \frac{38}{50} * 100 = 76\%$$

Figure 7 shows the true positive rate with respect to 50 different human biometric samples. The true positive rate of proposed DCD-WR framework is compared with two state-of-the-art methods [1], [2]. While increasing the biometrics samples, the true postive rate is increased in all three methods. But comparatively, proposed DCD-WR framework provides higher true positive rate than the LR-JSR and DCA. This is due to the application of deep learning template matching algorithm for accurate biometric recognition in DCD-WR framework. For each input biometric sample, the nonlinear activation function (sigmoid function) and hyperbolic tangent (tanh) functions, are determined using this algorithm. With the help of nonlinear activation function, the normalized image obtained and fusion template matching is carried out using hyperbolic tangent function. This helps to increases the true positive rate in DCD-WR framework. As a result, proposed DCD-WR framework increases the true positive rate by 7% compared to LR-JSR and 14% compared to DCA.

5.5. Impact of false acceptance rate

Lastly, the performance of false acceptance rate is evaluated using DCD-WR framework is and comparison is done with the existing Low-Rank and Joint Sparse Representations (LR-JSR) for Multi-Modal Recognition [1] and Discriminant Correlation Analysis (DCA) [2], respectively. The 500 input biometric samples are randomly considered for analysing the performance of false acceptance level. The sample calculations are given below,

Sample calculation

- **LR-JSR:** The 50 human biometric samples given as input, then the biometric samples incorrectly identified as 10. Therefore, false acceptance rate is calculated as,

$$FAR = \frac{10}{50} * 100 = 20\%$$

- **DCA:** The 50 human biometric samples given as input, then the biometric samples incorrectly identified as 13. Therefore, false acceptance rate is calculated as,

$$TPR = \frac{13}{50} * 100 = 26\%$$

- **Proposed (DCD-WR):** The 50 human biometric samples given as input, then the biometric samples incorrectly identified as 6. Therefore, false acceptance rate is calculated as,

$$TPR = \frac{6}{50} * 100 = 12\%$$

The comparison of false acceptance rate is depicted in Table 3 with respect to the different human biometric samples. The number of human biometric samples is taken as 50-500. With the increase in number of biometric samples, the false acceptance rate is increases in all three methods. From the table, comparatively proposed DCD-WR framework reduces the false acceptance rate while performing multimodal biometric recognition.

By considering 100 human biometric samples as input, the false acceptance rate of proposed DCD-WR framework is 12% whereas, the existing LR-JSR [1] provides 21% of false acceptance rate and DCA [2] provides 26% false acceptance rate. From the above discussion, the false acceptance rate of proposed DCD-WR framework is reduced than the state-of-the art methods.

The effective reduction on false acceptance rate is achieved with the help of contourlet transform in proposed DCD-WR framework. Contourlet transform performs pre-processing on multimodal biometric features where the irrelevant features are eliminated and choose more relevant features for biometric recognition. Moreover, deep learning template matching

Table 3. Tabulation of false acceptance rate.

Number of human biometric sample	False acceptance rate (%)		
	Proposed DCD-WR	Existing LR-JSR	Existing DCA
50	12	20	26
100	13	21	27
150	14	23	29
200	16	25	30
250	17	27	32
300	19	29	33
350	20	30	35
400	22	33	37
450	24	35	38
500	25	36	40

algorithm is applied to accurately match the biometric samples for multimodal biometric recognition. This in turns, the false acceptance rate is minimized in proposed DCD-WR framework. As a result, the DCD-WR framework reduces the false acceptance rate up to 35% and 45% when compared to existing LR-JSR [1] and DCA [2], respectively.

6. Conclusion

Multimodal biometric recognition is a complex task due to the development of sensing and multimodal communication technologies and the availability of high dimensional richer information about input biometric data. Therefore, there arises an urgent need to provide a mechanism for this high dimension and multimodalities and to propose an efficient framework to provide high recognition rate in the biometric recognition system. In this article, we provide Deep Contourlet Derivative Weighted Rank (DCD-WR) framework that can be employed as a biometric recognition method. Initially, the features were extracted using Local Derivative Ternary Pattern for face, fingerprint and iris images. With the extracted features, dominant attributes were stored in histogram form which resulted in the improvement of computational complexity involved in recognition. The evaluation of the template matching is performed finally using the Deep Learning Template Matching algorithm with the objective of improving the recognition rate. Through the experiments using real traces, we observed that our multimodal biometric recognition framework reduced computational time and computational complexity compared to the existing biometric recognition methods. In future, the redundant feature removal process is developed using boosting techniques for obtaining pre-processed features with minimum system error. Future work concentrates to use more biometric samples with different metrics to increase the scalability of the framework. In future work, DCD-WR framework is suggested to solve many problems like reducing recognition time, false acceptance level, accuracy and so on.

Disclosure statement

References

[1] Zhang H, Patel VM, Chellappa R. Low-rank and joint sparse representations for multi-modal recognition. IEEE Trans Image Process. 2017;26(10):1–13.

[2] Haghighat M, Abdel-Mottaleb M, Alhalabi W. Discriminant correlation analysis: real-time feature level fusion for multimodal biometric recognition. IEEE Trans Inf Forensics Secur. 2016;11(9):1–13.

[3] Gomez-Barrero M, Galbally J, Fierrez J. Efficient software attack to multimodal biometric systems and its application to face and iris fusion. Pattern Recognit Lett. 2013;3:243–253.

[4] Kanhangad V, Kumar A, Zhang D. Contactless and pose invariant biometric identification using hand surface. IEEE Trans Image Process. 2011;20(5):1415–1424.

[5] Choi H, Park H. A multimodal user authentication system using faces and gestures. Hindawi, BioMed Research International, No. 15, 2015. p. 1–8.

[6] Murillo-Escobar MA, Cruz-Hernández C, Abundiz-Péreza F, et al. A robust embedded biometric authentication system based on 4 fingerprint and chaotic encryption. Expert Syst Appl. 2015:2015:1–14.

[7] Long TB, Thai LH, Hanh T. Multimodal biometric person authentication using fingerprint. Face Features. Pacific Rim International Conference on Artificial Intelligence. 2012: 613–624.

[8] Yuan S, Zhang T, Zhou X, et al. An optical authentication system based on encryption technique and multimodal biometrics. Opt Laser Technol. 2013;54(2013): 120–127.

[9] Rajagopal G, Palaniswamy R. Performance evaluation of multimodal multifeature authentication system using KNN classification. Hindawi, The Scientific World Journal; 2015. p. 1–9.

[10] Saini N, Sinha A. Face and palmprint multimodal biometric systems using Gabor–wigner transform as feature extraction. Pattern Anal Appl. 2014;18(4):921–932.

[11] Saevanee H, Clarke N, Furnell S, et al. Continuous user authentication using multi-modal biometrics. Comput Secur. 2015;53:234–246.

[12] Lin WH, Wanga P, Tsai CF. Face recognition using support vector model classifier for user 4 authentication. Electron Commer Res Appl. 2016;18(2016):71–82.

[13] Ortega-Garcia J, Fierrez J, Alonso-Fernandez F, et al. The multi-scenario multi Environment BioSecure multimodal database (BMDB). IEEE Trans Pattern Anal Mach Intell. 2010;32(6):1097–1111.

[14] Khoury E, Shafey LE, McCool C, et al. Bi-modal biometric authentication on mobile phones in challenging conditions. Image Vis Comput. 2014;32(12):1147–1160.

[15] Le-qing Z, San-yuan Z. Multimodal biometric identification system based on finger geometry, knuckle print and palm print. Pattern Recognit Lett. 2010;31(12): 1641–1649.

[16] Galbally J, Marcel S, Fierrez J. Image quality assessment for Fake biometric Detection: application to iris, fingerprint, and face recognition. IEEE Trans Image Process. 2014;23(2):710–724.

[17] Choi JY, Neve WD, Plataniotis KN, et al. Collaborative face recognition for improved face annotation in personal photo collections shared on online social networks. IEEE Trans Multimedia. 2011;13(1):14–28.

[18] Poh N, Bourlai T, Kittler J. A multimodal biometric test Bed for quality-dependent, cost-sensitive and Client-specific score-level fusion algorithms. Pattern Recognit. 2010;43(3):1094–1105.

[19] Poh N, Windridge D, Mottl V, et al. Addressing missing values in kernel-based multimodal biometric fusion using neutral point substitution. IEEE Trans Inf Forensics Secur. 2010;5(3):461–469.

[20] Poh N, Kittler J. A unified framework for biometric Expert fusion incorporating quality measures. IEEE Transactions on Patterns Analysis and Machine Intelligence. 2012;34(1):3–18.

[21] Zhang S, Wang H, Huang W. Two-stage plant species recognition by local mean clustering and weighted sparse representation classification. Cluster Comput. 2017;20(2):1517–1525.

[22] Bakshi S, Sa PK, Wang H, et al. Fast periocular authentication in handheld devices with reduced phase intensive local pattern. Multimed Tools Appl. 2018;77(14):17595–17623.

[23] Zhang S, Wang H, Huang W, et al. Plant diseased leaf segmentation and recognition by fusion of superpixel. K-means and PHOG. Optik-International Journal for Light and Electron Optics. 2018;157:866–872.

[24] Zhang S, Wang H, Huang W, et al. Combining modified LBP and weighted SRC for palmprint recognition. Signal Image Video Process. 2018: 1–8.