A multimodal biometric identification system based on cascade advanced of fingerprint, fingervein and face images

El mehdi Cherrat¹, Rachid Alaoui², Hassane Bouzahir³
¹,³ISTI Laboratory, National School of Applied Sciences (ENSA), Ibn Zohr University, Morocco
²LRIT Laboratory, Faculty of Sciences, Mohammed V University, Morocco
³MUSICS TEAM, National Institute of Posts and Telecommunication (INPT), Morocco

ABSTRACT
In this paper, we present a multimodal biometric recognition system that combines fingerprint, fingervein and face images based on cascade advanced and decision level fusion. First, in fingerprint recognition system, the images are enhanced using gabor filter, binarized and passed to thinning method. Then, the minutiae points are extracted to identify that an individual is genuine or impostor. In fingervein recognition system, image processing is required using Linear Regression Line, Canny and local histogram equalization technique to improve better the quality of images. Next, the features are obtained using Histogram of Oriented Gradient (HOG). Moreover, the Convolutional Neural Networks (CNN) and the Local Binary Pattern (LBP) are applied to detect and extract the features of the face images, respectively. In addition, we proposed three different modes in our work. At the first, the person is identified when the recognition system of one single biometric modality is matched. At the second, the fusion is achieved at cascade decision level method based on AND rule when the recognition system of both biometric traits is validated. At the last mode, the fusion is accomplished at decision level method based on AND rule using three types of biometric. The simulation results have demonstrated that the proposed fusion algorithm increases the accuracy to 99.43% than the other system based on unimodal or bimodal characteristics.

Keywords: Biometric identification, Face recognition, Fingerprint recognition, Fingervein recognition, Fusion, HOG, LBP

1. INTRODUCTION
In recent years, the necessity for the biometric system has been rapidly increased. The biometric recognition is required to distinguish one individual from another using measurable morphological (such as fingerprint, face, iris, etc.) or behavioral (for example voice, signature, etc.) features. With these characteristics including being less susceptible to verification being stolen or forgotten. It is used for criminal identification, immigration and naturalization service, securing access to buildings or personal objects, supporting anonymous transactions, etc [1].

The most common biometric system is fingerprint recognition. It is considered an excellent biometric modality for identification or verification the person, especially in the latest smart phones and consumer devices [2, 3].

Compared to other biometric traits, the finger vein modality has achieved popularity in biometric recognition because of the variety advantages given by these systems for example, 1) the vein of each person are completely unique and different 2) it is identified as being less prone to modify with age and growth (3 the finger veins biometric is easily acquired using sensor capable of capturing or the NIR (Near-Infrared)
light source 4) the vein structure is hidden inside the skin. Thus, the possibility of spoof the human recognition system is very complex [4].

Face recognition is a biometric recognition technology based on human facial feature information for identification or verification. The algorithms using facial recognition are sensitive to variance in facial expressions and accessories, uncontrolled illumination, poses. In this regard, human and computer performance on facial identification is a research topic with both scientific research value and widely application prospects [5].

The general structure of biometric recognition system consists of four main steps. In the first one, the acquisition of biometric image is process of getting a digitalized image of a person using specific capturing device. In the second step, the pre-processing is allowed to improve overall quality of the captured image and to correct its orientation. After that, the region of interest is localized. It is the process of obtaining all important data needed for recognition. In the next step, the features information are extracted using different algorithms. In the last step, generally, the matching of the extracted characteristics is applied in order to perform the recognition of the person.

The multimodal biometrics combines two or more different biometric modalities and reduces certain limitations of systems based on one modality such as spoof attacks, non-universality, noise in sensed data, inter-class similarities and intra-class variations. Thus, the recognition system based on fusion of multibiometric is most recommended for significantly improving the system performance and reducing the error rate the identification or verification of the individual. This fusion can be applied at the sensor level, the feature extraction level, the matching-score level, rank level and at the decision level. The multimodal biometric systems are classified as multi-instance, multi-sensor, multi-algorithm, multi-modal and hybrid systems [6].

Many techniques have been proposed of the multimodal biometrics system. Ross et al presented different levels of fusion and score level fusion on the multimodal biometric system [6]. Singh.al proposed biometric recognition system based on face combining visible and thermal Infrared (IR) images at sensor level [7]. Connaughton et al. have been subjected a fusion of face and iris [8]. Ross et al. presented hand and face combined at feature level. Moreover, the experiments were applied in three different scenarios [9]. Different fusion techniques and normalization methods of fingerprint, hand geometry and face biometric sources are achieved by Jain.al [10]. Another multimodal biometric system based on multi-instance iris recognition system using a fusion of right iris and left iris for the same individual is studied by Wang et al. [11]. Jain et al. introduced a multimodal biometric system using face, fingerprint, and voice [12]. Yang et al. presented a multi-biometric system cancelable using fingerprint and finger-vein, which combines the minutia points of fingerprint and finger-vein image feature based on a feature-level of three fusion techniques [13]. The fusion multimodal biometric system based on fingerprint and finger-vein at score level using four score fusion approaches (min score, max score, simple sum, user weighting) and three score normalization techniques (min-max, z-score, hyperbolic tangent) is developed by Vishi et al. [14].

The rest of the paper is separated into four sections. Section 2 discusses the proposed algorithm. Experimental results have been analysed and discussed in Section 3. Finally, the conclusion is presented in the last section.

2. PROPOSED METHOD

In first level of our algorithm, fingerprint image is enhanced using gabor filter technique, binarized and passed to thinning algorithm [15]. Then, the features points are extracted using ridge ending and bifurcation uniformly namely minutiae. In the final step, the comparison of minutia information provided from the registered database, and the query fingerprint is presented to the matching. In the second level, the Linear Regression Line have been utilized to solve the orientation of misalignments of finger vein images. Next, the region of interest of image is obtained using Canny method [16]. After that, the histogram equalization [17] is applied to enhance the cropped finger vein image. Furthermore, the features extracted is based on Histogram of Oriented Gradient algorithm [18]. In the system based on face recognition, the Convolutional Neural Networks (CNN) [19] is applied to determine the size and position of the face in the image in order to extract informations due to Local Binary Pattern (LBP) method [20]. Finally, the scores provided are compared and matched with stored templates in the database. The fusion is applied at cascade advanced decision level. Therefore, the second level works only if the first level is not passed. Moreover, the third level is employed when the person is not identified using the first and the second level. The final fusion is accomplished at decision level method based on AND rule using three biometrics signatures. In this section, we detail the proposed technique which is illustrated in Figure 1. The details of each phase are represented in the following.

A multimodal biometric identification system based on cascade advanced of fingerprint... (El mehdi Cherrat)
Figure 1. Effects of selecting different switching under dynamic condition

2.1. Fingerprint Recognition System

2.1.1. Image Enhancement

To overcome the background noise, non-uniform illumination and low contrast of the fingerprint image captured, the preprocessing is important step for characteristic extraction and then the matching. The mean and variance are used to normalize and estimate the orientation of input image. After that, the frequency image is computed from which the region mask is provided using block classification of resulted image. Then, gabor filters applied to binarized image.

2.1.2. Minutiae Feature Extraction

The features points are extracted from fingerprint image such as ridge ending and bifurcation uniformly namely minutiae. Before extracting the minutiae, the binarization method is applied using block with size 3×3. This process is transformed the 8 bits gray image to 1 bit with 1 value for the valleys and 0 value for ridges based on a given threshold. Next, morphological technique processing (dilatation and erosion) is used as post-processing to achieve more compact blocks for reducing the noise region. Moreover, the thinning operation is applied to remove basically the redundant pixels until having a single pixel width \cite{Yang18}. Finally, the bifurcation and ending points are detected by computing black pixels of 8-directional nearest for each pixel point in fingerprint image. If the central pixel is black and has 3 black values nearest, then this pixel is a bifurcation. When the number of black nearest is just 1, the feature point is represented ending. The connection number (CN) for a given ridge can be represented in 5. Hence, the minutia characteristic extraction of fingerprint pattern is represented by the following parameters, 1) Type of the ridge, 2) \(x\)-coordinate, 3) \(y\)-coordinate and 4) \(\theta\)-orientation.

\[
CN = \frac{1}{2} \sum_{i=0}^{7} |P_i + P_{i+1}|
\]

(1)

where \(P_i\) is the pixel value at index \(i\) and \(P_8 = P_0\).

2.1.3. Features Matching

The minutiae feature extraction of fingerprint pattern is represented by the type of ridge, the spatial coordinates \(x\), \(y\) and orientation of minutiae points.
The Euclidian distance is used to find the number of matched two minutiae pairs. This distance $E_d$ is described as follows:

$$E_d(M_i, M_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$  \hfill (2)

where $M_i$ and $M_j$ are the extracted minutiae points pairs from the template in the enrolled database and the input query fingerprint image respectively.

The similarity score $S_{score}$ based on minutiae points between the queried and stored fingerprint images is calculated using (3).

$$S_{score} = \frac{N_m}{N_i N_j}$$  \hfill (3)

where $N_m$ is the total matching of $M_i$ and $M_j$, $N_i$ and $N_j$ are total number of $M_i$ and $M_j$ respectively.

$$N_m = \sum_{i=1}^{n} \sum_{j=1}^{n} M(M_i, M_j)$$  \hfill (4)

$$M(M_i, M_j) = \begin{cases} 1 & \text{if } (E_d(M_i, M_j) \leq r_0) \text{ and } (A_d(M_i, M_j) \leq \theta_0) \\ 0 & \text{Otherwise} \end{cases}$$  \hfill (5)

$$A_d(M_i, M_j) = \min(|\theta_i - \theta_j|, 360 - |\theta_i - \theta_j|)$$  \hfill (6)

where $r_0$ is the tolerance distance between $M_i$ and $M_j$, $A_d$ is the smaller direction difference between $M_i$ and $M_j$ than an angular tolerance $\theta_0$.

### 2.2. Finger vein Recognition System

#### 2.2.1. Orientation Correction

In this section, the obtained region of finger vein images using Canny method [16] is needed to determine that images are oriented correctly or not. The orientation corrected angle can affect to accurately extract feature extraction and matching. Therefore, the Linear Regression Line is applied to compute the estimated orientation angle $\theta$ represented in the Figure 2. First, all middle points are represented the line function of the finger vein image, which is defined in (7). Next, the orientation angle value $\theta$ is calculated by using (10). Finally, these images are considered normal, if orientation angle value is equal to 0, otherwise the finger vein image is not correctly oriented. The Figure 3 represents the results of the orientation corrected angle.

$$y = ax + b$$  \hfill (7)

$$a = \frac{\sum_{i=1}^{M} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{M} (x_i - \bar{x})^2}$$  \hfill (8)

$$\bar{x} = \frac{1}{M} \sum_{i=1}^{M} x_i , \quad \bar{y} = \frac{1}{M} \sum_{i=1}^{M} y_i$$  \hfill (9)

$$\theta = \begin{cases} \arctan(a) & \text{if } (a < 0) \\ \arctan(a) & \text{if } (a > 0) \\ 0 & \text{if } (a = 0) \end{cases}$$  \hfill (10)

where $M$ represent number of $x$ and $y$. The $a$ parameter is computed using (8).
2.2.2. ROI Detection

After contour detection and orientation correction, extraction of the rectangular region of interest (ROI) is the next step in the pretreatment process for finger vein images. This operation allows to locate and isolate the finger area of the image and remove the background of the image (undesired regions). Indeed, the width and the height of this region are obtained by the values of maximum and minimum abscissa of the profile of the finger. The result of Canny edge detector and ROI of finger vein image are shown in Figure 4.

2.2.3. Pre-Processing

To overcome the problems of light disturbance and noise in ROI, in order to obtain a uniform intensity distribution, we use the method of the local equalization of the histogram [17]. In fact, it makes it possible to increase the quality of the image and to improve the visibility of the veins in the sense that the performances of the following phase are maximized. The Figure 4(d) presents ROI contrast enhancement.

2.2.4. HOG Feature Extraction Method

HOG (Histogram of Oriented Gradients) descriptor has shown outstanding success in recognition system. HOG has been popular used as one of the better features to acquire local shape points or the edge [21]. For this advantage, this technique is applied in our algorithm for feature extraction in order to recognize the person. The HOG orientation of each cell, small connected areas, is separated. For better compensating the illumination, the normalized histogram is obtained by accumulating a measure of the local histogram gradient orientation over blocks based on the results to normalize each cell in the block. These histograms are combined to represent the HOG feature [18]. The process of extracting the HOG descriptor is illustrated in Figure 5.
2.2.5. Features Comparison

Before to compute the generated similarity score based on HOG features, Hamming distances $D_H$ is computed to match scores between the finger vein template stored in database and the input test template as calculated using (11). The similarity score of HOG $S_{hiscore}$ is given by (12).

$$D_H = \sum_{i=0}^{k} |F_{Ei} - F_{Ti}|$$

where $F_{Ei}$ and $F_{Ti}$ are the extracted HOG from the template in the enrolled database and the input query finger vein image respectively.

$$S_{hiscore} = \min(D_H)$$

2.3. Face Recognition System

2.3.1. Face Detection

The first task in face recognition system is face detection. This process is used to detect and locate faces in images. In our research, the mainly reason to use CNN [19] is obtained the higher performance accuracy of detection [22]. Table 1 shows the performance of CNN method and other traditional face detection techniques such Haar [23], LBP [24] and HOG based on SVM [25] in term of face detection accuracy rate for different faces databases.

| DataBases | Haar  | LBP   | HOG   | CNN  |
|-----------|-------|-------|-------|------|
| D[26]     | 78.56%| 72.63%| 92.12%| 100% |
| D[27]     | 98.64%| 89.80%| 93.90%| 99.90%|

2.3.2. LBP Feature Extraction Method

LBP (Local Binary Patterns) method is one of the best performing texture descriptors and widely used in various applications. It is proposed by T. Ojala et al. [28]. By definition, the LBP operator is robust against monotonic (lighting changes) gray scale transformations. For each pixel of an image, a binary code is produced to make a new matrix with the new value (binary to decimal value).

$$LBP_{pp}(N_c) = \sum_{p=0}^{r-1} g(N_p - N_c) \times 2^p$$

where $p$ is sampling points (e.g., $p = 0, 1, \ldots, 7$ for a window size 3x3), $r$ is radius for window size 3x3, $N_p$ is neighborhood pixels value and $N_c$ is center pixel value. The binary threshold function $g(x)$ is represented as follows:

$$g(x) = \begin{cases} 0 & \text{if} \ x < 0 \\ 1 & \text{if} \ x \geq 0 \end{cases}$$

In LBP method, the facial image is divided into local regions and LBP texture descriptors are extracted from each region independently. The descriptors are then concatenated to form a global description of the face.
2.3.3. Features Matching

The scores $S_{\text{score}}$ based on LBP method is obtained using SVM classifier in order to identify the face image. This technique of machine learning is used to separate the groups with the plane which will have the maximum margin (optimal hyperplane) [26].

2.3.4. Fusion

The decision-level fusion that integrates biometric signatures in a simple and straightforward way compared to other techniques. With these reasons, this strategy is employed in the proposed method to summarize the multiple results of the local decisions come from the three previously mentioned systems into a single overall result. In the unimodal cascade mode $\text{CD}_{\text{Multimodal}}$, the user is authentic once the decision of an identification system based on a single biometric mode is validated according to a given threshold. In the $\text{CD}_{\text{bimodal+ar}}$ bimodal and And Rule cascade mode, the decision is accepted as long as the person is identified by two biometric systems according to a condition. In multimodal and And Rule cascade mode $\text{CD}_{\text{Multimodal+ar}}$, the individual is an impostor if the decision of a single system is rejected. These three decisions are presented by the following:

$$\text{CD}_{\text{Multimodal}} = \begin{cases} \text{accepted, if } (D_{FP} \cup D_{FV} \cup D_{F}) = 1 \\ \text{rejected, else} \end{cases}$$

(15)

$$\text{CD}_{\text{bimodal+ar}} = \begin{cases} \text{accepted, if } (D_{FP} \cap (D_{FV} \cup D_{F}) \cup (D_{FV} \cap D_{F})) = 1 \\ \text{rejected, else} \end{cases}$$

(16)

$$\text{CD}_{\text{Multimodal+ar}} = \begin{cases} \text{accepted, if } (D_{FP} \cap D_{FV} \cap D_{F}) = 1 \\ \text{rejected, else} \end{cases}$$

(17)

where $D_{FP}$, $D_{FV}$, $D_{F}$ correspond to the decision taken regarding the person to be recognized in the system of fingerprints, fingerveins and faces respectively.

3. RESULTS AND ANALYSIS

The experimental operation platform in this study is described as follows: the host configuration: CPU Intel Core2 Duo at 2.00 GHz, RAM 3.00 GB, runtime environment: Microsoft Visual Studio C++ 2013 with OpenCV and Dlib library. In order to validate the proposed algorithm, the results have been tested on the public Fingerprint Verification Competition 2004 dataset [29], the VERA Fingervein Database [30] and The AR face database [27]. The performance measure is accuracy rate as defined by (15).

$$\text{Accuracy} = \frac{TP + TN}{\text{TotalNumAcc}}$$

(18)

where $TN$ (True Negative Rate) is the probability of authorized users that are recognized correctly over the total number tested, $TP$ (True Positive Rate) describes the probability of authorized users that are not recognized over the total number tested and $\text{TotalNumAcc}$ is the total number access.

Table 2 shows the performance of accuracy rate based on single biometric system using fingerprint, fingervein, face images and the cascaded multimodal recognition biometric system using two and three biometric traits. In comparison with single biometric system, our proposed algorithm especially with the cascaded multimodal biometric system using fingerprint, fingervein and face images shows superior performance in terms of accuracy rate with 99.43% with where fingerprint using minutiae points, fingervein using HOG, Face using LBP, bimodal, multimodal cascaded mode and And Rule give 96.40%, 97.34%, 90.86%, 99.03% and 96.28% respectively. We can conclude from these results that the cascaded multimodal recognition biometric system using fingerprint, fingervein and face images leads to an improvement in recognition biometric system performance.

| Algorithms                      | Accuracy Rate |
|---------------------------------|---------------|
| Fingerprint using Minutiae      | 96.40%        |
| Fingervein using HOG            | 97.34%        |
| Face using LBP                  | 90.86%        |
| Cascaded Multimodal             | 99.43%        |
| Cascaded bimodal and And Rule   | 99.03%        |
| Cascaded Multimodal and And Rule| 96.28%        |

Table 2. The Accuracy Rate for Different Recognition Biometric System Results
4. CONCLUSION
This article focuses on the proposition of the cascading merge identification system at the decision level based on three biometric signatures, fingerprints, finger veins and faces, to provide accurate recognition of the person. At the fingerprint identification level, the image is improved according to the Gabor filter algorithm, binarization and thinning technique in order to extract and compare the minutia points. At the finger vein identification level, segmentation based on the Canny method, orientation correction, ROI detection and local histogram equalization are applied to improve the quality of the images. In order to better identify the relevant information, the HOG approach is applied in the extraction module. At the face identification level, the LBP method is used to construct the feature vector from faces detected by CNN for matching purposes. The identification process for each system is accepted in view of the scores being higher than a given threshold. The unimodal cascade mode is used when the decision of a single system is accepted. However, the bimodal cascade mode consists of having at least the favorable decision of two systems. In the multimodal cascade mode, the person will be considered an impostor once the decision of a single system is rejected. Our experimental results show that this proposed work ensures promising identification performance, especially on the cascade decision fusion of these three modalities.

REFERENCES
[1] E.M. Cherrat, et al., "High density salt-and-pepper noise suppression using adaptive dual threshold decision based algorithm in fingerprint images," IEEE Intelligent Systems and Computer Vision (ISCV), pp. 1-4, 2017.
[2] E.M. Cherrat, et al., "Improving of fingerprint segmentation images based on K-means and DBSCAN clustering," International Journal of Electrical and Computer Engineering (IJEC), vol. 9, p. 2088-8708, 2019.
[3] S. R. Borra, et al., "An Efficient Fingerprint Identification using Neural Network and BAT Algorithm," International Journal of Electrical and Computer Engineering (IJEC), vol. 8, no 2, p. 2088-8708, 2018.
[4] C. H. Hsia, "New Verification Strategy for Finger-Vein Recognition System," IEEE Sensors Journal, vol. 18, no 2, pp. 790-797, 2018.
[5] D. N. Z. A. Jesemi, et al., "The analysis of facial feature deformation using optical flow algorithm," Indonesian Journal of Electrical Engineering and Computer Science (IJEECS), vol. 15, no 2, p. 758-768, 2019.
[6] A. Ross and J. Anil, "Biometric sensor interoperability: A case study in fingerprints," In ECCV Workshop BioAW, pp. 134-145, 2004.
[7] R. Singh et al., "Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition," Pattern Recognition, vol. 41, no 3, p. 880-893, 2008.
[8] R. Connaughton, et al., "Fusion of face and iris biometrics," Handbook of Iris Recognition Springer, pp. 219-237, 2013.
[9] A. Ross and R. Govindarajan, "Feature level fusion of hand and face biometrics," Biometric Technology for Human Identification II. International Society for Optics and Photonics, p. 196-205, 2005.
[10] A. Jain et al., "Score normalization in multimodal biometric systems," Pattern recognition, vol. 38, no 12, pp. 2270-2285, 2005.
[11] P. Wang, "Pattern recognition, machine intelligence and biometrics," Springer Berlin Heidelberg, 2011.
[12] A. Jain et al., "A multimodal biometric system using fingerprint, face and speech," 2nd Int'l Conf. AVBPA, vol. 10, 1999.
[13] W. Yang, et al., "A fingerprint and finger-vein based cancelable multi-biometric system," Pattern Recognition, vol. 78, pp. 242-251, 2018.
[14] K. Vishi and S. Y. Yayilgan, "Multimodal biometric authentication using fingerprint and iris recognition in identity management," Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing. IEEE, p. 334-341, 2013.
[15] L. Hong et al., "Fingerprint image enhancement: algorithm and performance evaluation," IEEE transactions on pattern analysis and machine intelligence, vol. 20, no 8, p. 777-789, 1998.
[16] J. Canny, "A computational approach to edge detection." Readings in computer vision. Morgan Kaufmann, pp. 184-203, 1987.
[17] J. Y. Kim, et al., "An advanced contrast enhancement using partially overlapped sub-block histogram equalization," IEEE transactions on circuits and systems for video technology, vol. 11, no 4, pp. 475-484, 2001.
[18] D. Navneet, and B. Triggs, "Histograms of oriented gradients for human detection," IEEE Computer Society International Conference on computer vision and Pattern Recognition, pp. 886-893, 2005.
[19] D. E. King, "Dlib-ml:A machine learning toolkit," Journal of Machine Learning Research, vol. 10, no Jul, pp. 1755-1758, 2009.
[20] X. Wang et al., "An HOG-LBP human detector with partial occlusion handling," IEEE 12th international conference on computer vision, p. 32-39, 2009.
[21] Enrique V. Carrera, et al., "A Finger-vein Biometric System Based on Textural Features," Journal of Power Electronics, pp. 367-375, 2018.
[22] K. Zhang, et al., "Joint face detection and alignment using multitask cascaded convolutional networks," IEEE Signal Processing Letters, vol. 23, no 10, p. 1499-1503, 2016.
[23] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features," null. IEEE, pp. 511, 2001.

A multimodal biometric identification system based on cascade advanced of fingerprint... (El mehdi Cherrat)
[24] L. Zhang, et al., “Face detection based on multi-block lbp representation,” International Conference on Biometrics. Springer, Berlin, Heidelberg, pp. 11-18, 2007.

[25] D. E. King, "Max-margin object detection," arXiv preprint, arXiv:1502.00046, 2015.

[26] P.J. Phillips et al., "The FERET database and evaluation procedure for face-recognition algorithms," Image and vision computing, vol. 16, no 5, p. 295-306, 1998.

[27] A.M. Martinez, "The AR face database," CVC Technical Report24, 1998.

[28] T. Ojala et al., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, no 7, p. 971-987, 2002.

[29] D. Maio et al., "Third fingerprint verification competition," International Conference on Biometric Authentication. Springer, Berlin, Heidelberg, p. 1-7, 2004.

[30] P. Tome et al., “On the vulnerability of finger vein recognition to spoofing,” International Conference of the Biometrics Special Interest Group (BIOSIG), p. 1-10, 2014.