

# SCORE LEVEL AND RANK LEVEL FUSION FOR KINECT-BASED MULTI-MODAL BIOMETRIC SYSTEM

Md Wasiur Rahman, Fatema Tuz Zohra, Marina L. Gavrilova\*

*Department of Computer Science  
University of Calgary  
Calgary, Alberta, Canada*

*\*E-mail: mgavriilo@ucalgary.ca*

*Submitted: 6th February 2018; Accepted: 27th September 2018*

## Abstract

Computational intelligence firmly made its way into the areas of consumer applications, banking, education, social networks, and security. Among all the applications, biometric systems play a significant role in ensuring an uncompromised and secure access to resources and facilities. This article presents a first multimodal biometric system that combines KINECT gait modality with KINECT face modality utilizing the rank level and the score level fusion. For the KINECT gait modality, a new approach is proposed based on the skeletal information processing. The gait cycle is calculated using three consecutive local minima computed for the distance between left and right ankles. The feature distance vectors are calculated for each person's gait cycle, which allows extracting the biometric features such as the mean and the variance of the feature distance vector. For Kinect face recognition, a novel method based on HOG features has been developed. Then, K-nearest neighbors feature matching algorithm is applied as feature classification for both gait and face biometrics. Two fusion algorithms are implemented. The combination of Borda count and logistic regression approaches are used in the rank level fusion. The weighted sum method is used for score level fusion. The recognition accuracy obtained for multi-modal biometric recognition system tested on KINECT Gait and KINECT Eurocom Face datasets is 93.33% for Borda count rank level fusion, 96.67% for logistic regression rank-level fusion and 96.6% for score level fusion.

**Keywords:** Lane-Emden equations, simulated annealing, legendre polynomials, neural network

## 1 Introduction

Computational intelligence firmly made its way into the areas of consumer applications, banking, education, commerce and social security [1, 4]. The biometric domain of research is one of the key areas where the benefits of using machine learning and computational intelligence methods cannot be underestimated [5]. Biometric systems rely on

user's physiological and behavioral characteristics for identification and verification purposes [1].

The multimodal biometric system emerged as a more reliable and secure biometric system, which utilizes two or more identifiers and thus overcomes the limitations of the unimodal system. Multimodal systems mitigate the inter-class similarity between two users of a system, increase the degree of freedom and reduce the spoof attacks [2]. They also

ensure sufficient population coverage, which mitigates the non-universality problem of unimodal systems. As a result, deployment of the multimodal biometric system in real-world identification systems is an increasing trend.

Traditionally, researchers used a standard video camera to collect the data [1]. But this type of data requires heavy pre-processing as well as sufficiently large storage. In addition, feature extraction from video sequences is computationally demanding and costly. The invention of Microsoft KINECT is changing the landscape of the traditional gait and face recognition biometric domains [7]. The KINECT device can provide 3D skeletal joint information for the moving human body, and display information in grayscale, RGB and depth image formats. The availability of API's for data process make the KINECT a convenient tool for feature extraction [3]. In this work, we have proposed a novel multimodal system using Microsoft KINECT face and gait datasets and compared the performance of two information fusion approaches: rank level and score level fusion.

Fusion scheme plays a vital role in the multimodal biometric system. A fusion is a technique that combines the information which is presented by multiple domain experts. Thus, the goal of the fusion is to determine the best set of experts for a given problem domain [2]. There are three traditional fusion approaches in the multimodal biometric system: match score level fusion, rank level fusion and decision -level fusion. In match score fusion, multiple classifiers are used to set the match scores that are fused to form a final scalar score. In decision level fusion, the final decision of acceptance and rejection is made using the majority voting type scheme from the outcomes of the individual biometric identifiers [6]. The rank level fusion computes the final rank of a user based on the ranks obtained from the individual classifiers [7]. Various techniques such as Borda count and logistic regression are used to make the final decision based on the rank of each identity [7]. In this work, we have considered both match score and rank level fusion using average sum, Borda count and logistic regression techniques. These techniques are widely used in the multimodal biometric system because they provide a high accuracy and are computationally not expensive. This research makes a number of im-

portant new contributions to the biometric domain. First, it uses KINECT data for both modules which provides easy data access using a single biometric device. To the best of our knowledge, there are no multimodal systems based on two KINECT modalities developed previously. Another contribution is that for Gait modality, the recognition is based on the skeletal information, while previous methods used silhouette data which lacks discriminability due to the person's outfit, clothes, and occlusions and objects a person may carry. A recently developed method based on all possible joint distances is utilized for gait recognition [8]. For KINECT face recognition, a novel method based on HOG features has been proposed in this work. Advantages of using HOG features lie in their ability to extract pertinent information from the gradient intensity of the facial image. This work demonstrates the high suitability of HOG features for KINECT facial data. Finally, the multimodal system performance using rank level and match score level fusions are implemented and evaluated on the virtual KINECT gait and face multimodal dataset. This work is an extended version of a conference article presented at 2017 IEEE Symposium Series on Computational Intelligence (SSCI'17) conference and invited as a submission in an extended form [10]. Specifically, a new method for information fusion based on score level is implemented and compared to previously reported results in this extended version.

## 2 Literature Review

There is ample research conducted over the past decade on multimodal biometric systems using a combination of different modalities at different fusion techniques [6, 7]. Researchers had started to study the fusion of face and gait biometrics in 2002 [10]. Shakhnarovich et al. [10] proposed a multimodal based system using the the frontal image and person silhouette for 26 subjects. They applied cross-fold fusion method for various features such as min, max, mean and product. Their system reached 89% accuracy by using the product-rule fusion scheme. After a couple of years, Kale et al. [11] proposed a system for 30 subjects using the similar feature extraction method but utilizing different fusion techniques: hierarchal and holistic. Ranganath et al. [12] applied chaos and Lyapunov

exponent in their system in 2006. In the same year, Zhou et al. [13] developed gait and face fusion from a distant system. The authors utilized gait energy image and high-resolution face image and applied principal component analysis (PCA) and multiple discriminant analysis (MDA) for feature extraction for 46 subjects. The outcome of their analysis showed that the integrated face and gait features carry the most relevant information over an individual modality. They obtained the accuracy of 91.3% when combining the two modalities. Zhou et al. [14] extended their work by utilizing Enhanced Side Face Image (ESFI) instead of a high-resolution face image. After using fusion technique, they concluded that cohesive information from side gait and frontal face is more effective for person recognition in video. They obtained 91.1% recognition using feature level fusion and 88.9% recognition using match score level fusion. Geng et al. [15] conducted a different research on gait and face fusion. They showed that view angle and the person-to-camera distance are the main aspects that may affect the relationship between face and gait in the fusion. They utilized min, max, mean technique to achieve the 90% accuracy of their adaptive fusion multimodal system.

In 2011, Emdad et al. [16] proposed to use principal component analysis (PCA) and linear discriminant analysis (LDA). They utilized hierarchical and holistic fusion and observed that a combination of PCA-LDA ensured better recognition rate than PCA features alone. Later, Hofmann et al. [17] proposed to utilize a low-resolution face image and obtained an accuracy reaching 92%. Almahmoud et al. [18] represented face through Active Lines among Face Landmark Points (ALFLP) feature vector and gait as Active Horizontal Levels (AHL) feature vector. Their method showed that the integrated face and gait features result in 96% accuracy using feature fusion. Most recently, Xing et al. [19] compared match score and feature level score and concluded that the latter provides a slight edge. Finally, interesting applications of the face and gait fusion for ethnicity classification [20] and gender identification [21] were recently considered.

In all the previous works, researchers considered the gait silhouette-based energy image. But this approach has two serious deficiencies. One is that silhouette cannot retain the subject's dynamic

information, and another is that silhouette is highly affected by a subject wearing heavy clothes, carrying objects, occlusions etc. With the invention of KINECT, that provides the 3D skeletal information in real-time, these problems can be solved in an efficient way. Thus, we propose to use the model-based gait recognition approach based on the 3D skeletal information. To the best of our knowledge, no work has been done using the gait 3D skeletal information fused with the face. In this work, we propose a novel multimodal biometric system that combines 3D skeletal gait information with KINECT face using rank level and match-score level fusions.

### 3 Proposed Methodology

The purpose of this work is to present a novel multimodal fusion biometric system for KINECT gait and face images. To achieve this goal, features of both gait and face modalities are extracted separately and entered into the system database at an enrollment phase. Then, during an identification phase, a feature matching algorithm based on KNN is utilized. The popular rank level fusion methods are Borda count and logistic regression. They have been used to investigate the overall accuracy of the multimodal system. In addition, match score level fusion using an average sum method was investigated. Figure 1 represents the block diagram of the proposed methodology.

#### 3.1 Gait recognition

A novelty of this work is that for Gait modality, the recognition is based on skeletal information, while previous methods used silhouette data. This limited accuracy due to the lack of discrimination in the silhouette data. We utilize a recently proposed method based on all possible joint distances for gait recognition [8]. In this work, we have used 3D skeletal data from KINECT v1 sensor for 30 persons. Each 3D skeleton consists of 20 joints. Gait cycle is detected for each person in the feature extraction process.

Gait cycle is calculated in order to extract the important features for person identification. A complete gait cycle consists of three steps as rest standing position-to-right foot forward-to rest-to left foot forward-to rest and vice versa [22]. So, the hori-

zontal distance between left and right ankle is calculated for each recorded frame. Then a moving average filter is used to smooth this horizontal distance vector. Finally, a gait cycle is calculated by detecting three consecutive local minima.

$$d_i = (x_{leftankle} - x_{rightankle}), \quad (1)$$

here,  $i = [1, N]$ , where  $N =$  the total number of frames in a video of a person walking.

Features of gait recognition can be extracted in various ways. In this work, we compute a distance feature vector. This feature vector calculates the distance from a joint to all other adjacent joints. The distance between two skeletal joints  $j_1(x_1, y_1, z_1)$  and  $j_2(x_2, y_2, z_2)$  can be defined as

$$d_{j_1 j_2} = \sqrt{|x_1 - x_2|^2 + |y_1 - y_2|^2 + |z_1 - z_2|^2}. \quad (2)$$

The feature distance vector for each joint is calculated with respect to others 19 joints for each person in a gait cycle. So, the total distance joint for each frame is 380. Then features as mean, variance is calculated for each joint for using as an input of the KNN classifier. We have considered 180 joints for each feature as the input to KNN. Note that the distance from one joint to another joint is:  $d(a, b) = d(b, a)$ . Here,  $a$  and  $b$  define the joints, and  $d()$  defines the distance function.

### 3.2 Face recognition

For KINECT face recognition, a novel method based on Histogram of Oriented Gradient (HOG) features has been developed. In this work, we have utilized KINECT face dataset for 30 persons. First, the face of each person is detected and then detected face is cropped. Then, the cropped face is divided into four equal parts. Hog feature is extracted from each equal part of each person's face. The advantage of using HOG features lies in their ability to extract pertinent information from the gradient intensity of the facial image. We have used vision cascade object detector to detect the face and then the detected face has been divided into four equal parts to increase the number of samples of the HOG feature for better recognition rate. Then each part is assigned for the feature extraction purpose. The face is divided so that feature vector can easily detect the face of a correct person when identification is done using the classification technique.

The HOG feature image represents the distribution of intensity gradients or edge directions [23]. The image is divided into a small regions, known as cells, where each cell contains pixel. Each cell is placed into an angular bin according to the gradient orientation. A weighted gradient is assigned to its corresponding angular bin and gradient intensity. Then histogram is plotted using gradient direction. In this work, we have utilized the HOG feature because we are using two scenarios of face dataset: one is obtained under light illumination and another is under the natural condition. A HOG feature represents the image according to its gradient intensity, which will differ for two different conditions. Our results demonstrate that using HOG features allows achieving a high accuracy even under less than ideal illumination conditions.

### 3.3 Fusion Techniques

Finally, the multimodal system is developed using two well performing post-matching fusion techniques. We first discuss the purpose of the using fusion method in the multi-modal system and then proceed to describe no example rank and match score fusion methods.

Previous research has demonstrated that integrating results from more than one biometric modality leads, in general, to a higher recognition rate [7]. The resilience to spoof attacks and ability to handle noisy data, intra-class variability and inter-class similarities have led to multi-modal systems being accepted as an industry standard. However, the type of biometric information that need to be integrated, the choice of information fusion approach, cost and benefit analysis needs to be conducted in order to guarantee a success in a reallife scenario. In order to gain a deeper understanding of those choices, in this paper we compare for the first time two common fusion approaches on example of KINECT sensor gait and face recognition.

Depending on the type of multimodal biometric information that needs to be fused, the various levels of multimodal biometric information fusion can be classified into two broad categories: fusion before matching and fusion after matching. Fusion before matching category means raw data or raw features are combined before they are matched again a sample to find the correct identity. They are usually called *sensor level* and *feature level fusion*, respec-



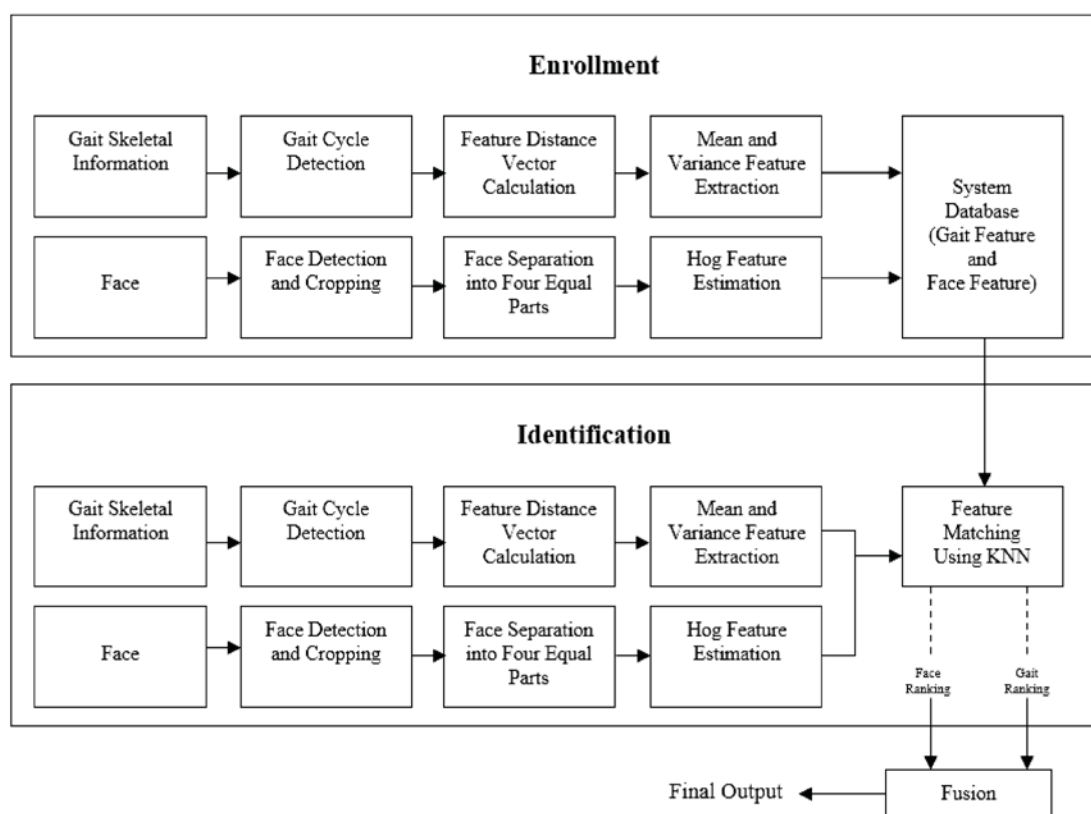


Figure 1. Block diagram of the proposed multi-modal identification system

tively. Fusion after matching contains **match score level fusion**, **rank level fusion** and *decision level fusion*, where individual biometric modalities of a sample are being matched to those in the individual biometric databases [1]. In our case, these are KINECT face and KINECT gait databases.

In the rank level fusion, two or more individual matches' outputs produces the ranking of the "candidates" in the template database sorted in a decreasing order of match scores. The system is expected to assign a higher rank to a biometric template from the database that is more similar to the given sample. In rank level fusion, rank is assigned to each database sample that is close person after obtaining the identification results from each individual identifier. Lowest rank (i.e. rank 1) represents the highest accuracy (closest possible matches). Finally, the ranked output of two individual biometrics is consolidated by using two popular methods: the Borda count and the logistic regression methods. We implemented both methods. We choose 0.3 and 0.1 as weights for face and gait respectively in the logistic regression method. The weight of the face is selected from the results established in the literature and is set to 0.3 [7]. How-

ever, no previous research has been done using the gait skeletal information for rank level fusion. So, we set the weight for KINECT gait at 0.1 as this modality has the higher overall stability than the face modality.

Figure 2 shows an example of rank level fusion using Borda count and logistic regression methods. Here, the rank of Person 1 is one for each face and gait modality. For the Borda count, the ranks obtained by two individual for face and gait are added and then divided by 2 (the number of matches). For the logistic regression, weighted rank is computed by multiplying the obtained ranks by predetermined weights. Finally, personal identities are reordered according to new ranks computed individually by Borda count and logistic regression methods.

As another contribution, we propose a lightly enhanced procedure that uses not one of, but both Borda count and logistic regression method outcomes. If two of those methods point to the same identity at the same rank, it stays. However, if there is a tie between ranks (see Figure 2 Borda Count Person 2 and person 3 fused scores are identical at 2.5), traditional systems use "coin toss" to ran-

domly break the tie, where we use the results of the second method logistic regression in this case, to break the tie for additional accuracy.

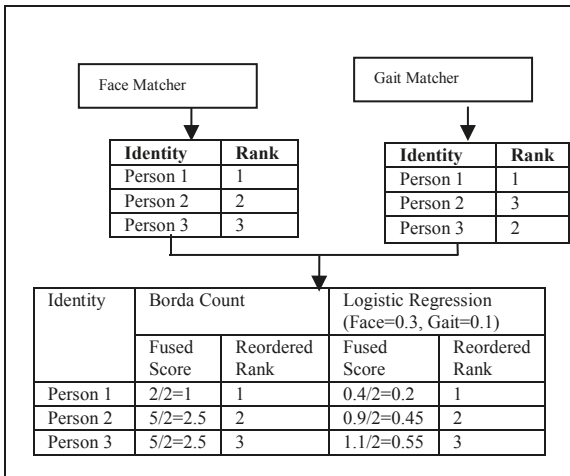


Figure 2. Example of rank level fusion

We now describe the system implemented using score level fusion. For obtaining a single matching score, this fusion method applies arithmetic operations, such as addition, subtraction, or median, to different matching scores. As an example, the match scores generated by two different matches for the face and gait may be combined by the sum rule in order to obtain a new match score which is then used to make the final decision. In our case, we utilize weighted match score sum rule, to obtain the fused score (see Figure 3). The match score fusion scheme combined the matching scores from two different classifiers. The match score defines the similarity between training and testing template. After obtaining the identification results from each individual biometric sample, matching score has been computed for each person’s sample for each of the modalities (face and gait). Then all the matching scores are normalized in the 1 to 100 range. Finally, we calculate the weighted sum matching score for each person from different modalities. The final decision on the matched identity in Figure 3 is Person 2 with the highest fused score of 89.5. Thus, Person 2 is the system output as the most likely person whose biometric templates match the sample provided to the multi-modal system.

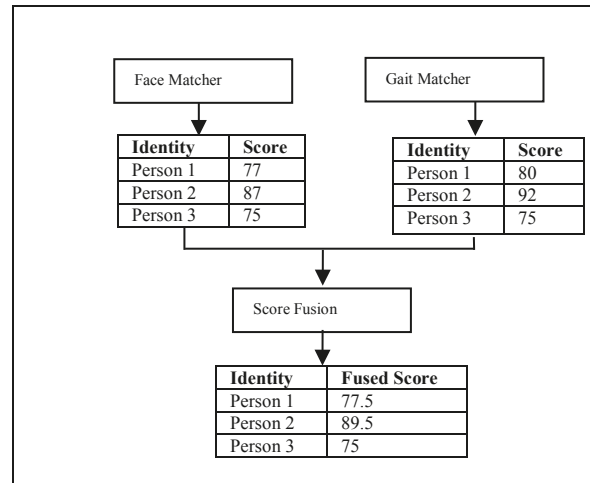


Figure 3. Example of match score level fusion

## 4 Experimental Results

We have considered KINECT v1 skeletal gait database [24]. The database consists of 30 persons among them 15 are males and 15 are females. The KINECT sensor was placed 1.70m above the ground. There are five video sequences for each person. So, total of 150 video sequences was stored in the database. For each sequence, a person was walking in a straight line with normal walking speed. Each gait sequence consists of 55 to 120 frames. Each person skeleton consists of 20 skeletal joints. All the skeletal joints of the dataset are shown in Figure 4. The dataset is publicly available from [24].

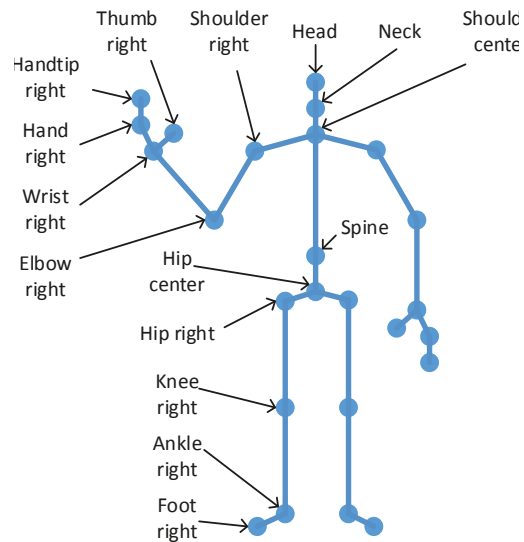


Figure 4. Example of match score level fusion

For the KINECT face modality, we used EUROCOM Kinect face dataset [25]. The dataset contains 52 subjects: 38 males and 14 females. The images were captured in two sessions at different time periods. The database contains three different sources of information: depth bitmap image, RGB image, and 3D objects. We have considered RGB face images for high illumination and natural conditions for both sessions. Each image consists of 256x256 pixels. Figure 5 and 6 shows the example of Kinect RGB face image with light on and natural condition. The training database plays a vital role in gaining better recognition performance from a biometric system. Due to the lack of true databases, researchers resolve to the use of virtual databases [2, 7]. It is a common practice in the multimodal domain to use the virtual database, which combines individual modalities from one or more true biometric databases, as long as the modalities are independent. In this work, the multimodal biometric system is a virtual database, where one person's identifier, i.e. gait, is combined with another person's identifier, i.e. face. The face dataset was reduced to 30 individuals to match the number of subjects in the gait dataset [25].

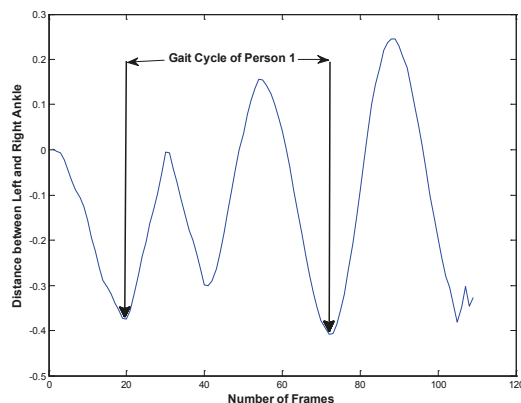


**Figure 5.** RGB face dataset with 'Light on' condition



**Figure 6.** RGB face dataset with 'Natural' condition

Figure 7 demonstrates the distance between the left and right ankle after passing moving average filter for person 1. The X-axis defines the number of the frame and the Y-axis represents the distance between the left and right ankle. After detecting three consecutive local minima in Figure 6, the result illustrates that the gait cycle lies between frame number 19 to 71.



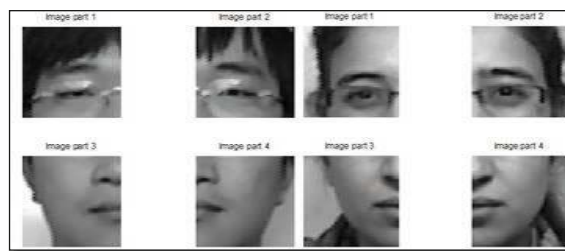
**Figure 7.** Gait cycle detection for Person 1

After obtaining the gait cycle, we calculate distance feature vector for each person in a gait cycle. Then mean and variance are calculated in the distance feature vector for each joint. Thus, we obtain  $(20 \times 19)$  or 380 mean and variance samples for each subject. Figure 8(a) shows the detection of the face of person 1 and person 16 using cascade object detector. After detecting the face, we crop the face to extract the feature which is shown in Figure 8(b).



(a) (b)

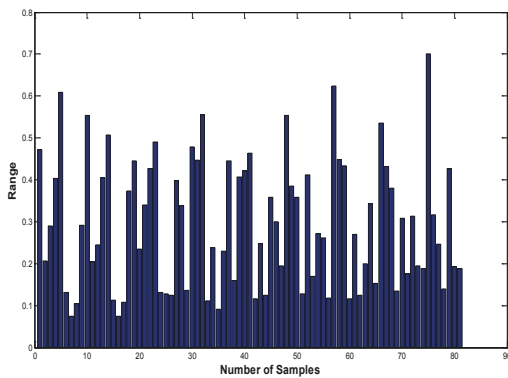
**Figure 8.** (a) Face Detection (b) Face Crop



**Figure 9.** Face subdivision into four equal parts

Processing of a face for each person involves dividing the image into four equal parts, as discussed earlier. Figure 9 represents the four equal parts of subject 1 and subject 16. HOG feature is extracted for each part individually. Figure 10 represents the histogram of person 1 for part 1. The X-axis defines the number of pixel sample and the Y-axis represents the value of each pixel in HOG transform. After obtaining the features for both face and gait, we have implemented KNN algorithm

for feature classification. We obtained the accuracy of 70% for feature variance ( $f_v$ ), 83.33% for feature mean ( $f_m$ ) and 93.33% for combined feature mean-variance ( $f_c$ ) respectively, using gait. We obtained 70% recognition for natural face condition ( $f_n$ ), 83.33% for light on ( $f_l$ ) face and 90% for combined light on and natural face conditions ( $f_c$ ). Table 1 demonstrates the accuracy of gait and face recognition using various features.



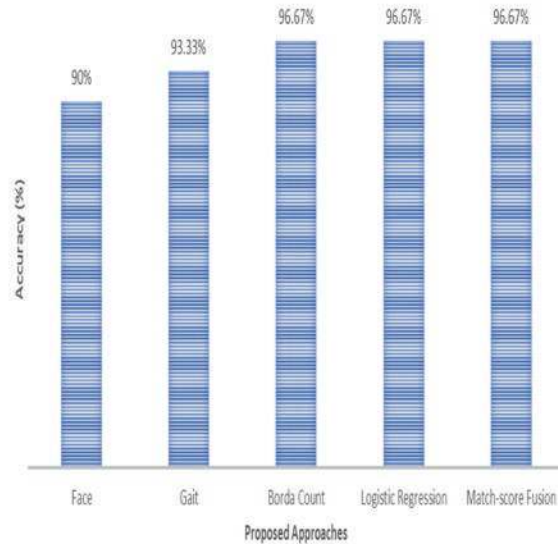
**Figure 10.** Hog Feature Histogram of Part 1 of Subject 1

**Table 1.** Classification of gait and face using different types of features and their combination

Method	Database	Type of Features	Accuracy (%)
Gait Skeletal Information	UPCV Kinect V1	Mean	83.33%
		Variance	70%
		Combined	93.33%
RGB Face Image	Eurocom Kinect Face	Light on	83.33%
		Natural	70%
		Combined	90%

From Table 1, we observe that the combined features provide the highest accuracy. Thus, we have used the combined features in the multimodal system. We have implemented the Borda count, logistic regression and average sum methods in the multimodal system. Rank is assigned by performing feature matching of the combined features of individual identifiers from the enrollment phase with the features of the identification phase. The lowest rank is assigned to the person with the highest classification accuracy. Figure 11 represents the overall accuracy of the developed multi-modal system. Unimodal techniques provide the accuracy of 90% for face and 93.33% for gait. This result is further improved to 93.33% using the Borda

count, 96.67% using the logistic regression and 96.67% using match-score fusion multimodal biometric system. Table 2 represents the summary of the recognition rates obtained by the most recent works utilizing gait and face biometrics in a multimodal system. It can be seen that the proposed method surpasses recognition results of all of the other multimodal systems reviewed.



**Figure 11.** Overall accuracy of the developed approaches

**Table 2.** Classification of gait and face using different types of features and their combination

Method	Year	Different form of face and gait modalities	Accuracy (%)
Hofmann et al. [17]	2012	Low-resolution face image and gait silhouette	92%
Almahmmod et al. [18]	2012	Front face and side gait	96%
Xing et al. [27]	2015	Face and gait	96%
Proposed Rank-level Approach	2017	Kinect gait and face	96.67%
Proposed Match-score Level Approach	2018	Kinect gait and face	96.67%



## 5 Conclusion

In this work, we presented for the first time the multimodal biometric system using KINECT face and gait modalities. By using a novel 3D KINECT joint based feature extraction method and proposing a HOG feature-based KINECT face recognition method, the accuracy of individual modalities was improved to over 90%. The system reaches the accuracy of 93.33% for combined features using K-nearest neighbors (KNN) feature classification algorithm for gait recognition. We have obtained the highest accuracy 90% for combined 'light on' and natural face image condition using the KNN algorithm for face. Implementing rank level and match score level fusions by combining the features of gait and face using Borda count, logistic regression and average sum approaches resulted in achieving even higher accuracy of 93.33% for Borda count, 96.67% for logistic regression and 96.67% for match score fusion. The findings of this work can be used as a foundation for future research in KINECT gait and face based real-time identification systems.

## Acknowledgements

We would like to acknowledge NSERC Discovery Grant RT731064, as well as MITACS and AITF for partial funding of this project.

## References

- [1] Gavrilova M., Monwar M., Multimodal Biometrics and Intelligent Image Processing for Security Systems, Hardcover, IGI Global, 350 pages, 2012
- [2] Ross A., Nandakumar K., Jain A., Handbook of Multibiometrics, New York: Springer-Verlag, 2006
- [3] Kastaniotis D. et.al. A Framework for Gait-based Recognition Using Kinect, Pattern Recognition, Elsevier, Vol. 68(2), pp. 327-335, 2015
- [4] Leszek Rutkowski, Computational Intelligence: Methods and Techniques, 1st Ed, book, Springer, 2008
- [5] Y Wang et.al., Cognitive Intelligence: Deep Learning, Thinking, and Reasoning by Brain-Inspired Systems, IJCINI, vol. 10, no. 4, pp1-20, 2016
- [6] Paul P.P., Gavrilova M., Alhadj R., Decision Fusion for Multimodal Biometrics Using Social Network Analysis, IEEE Trans. On Systems, Man, and Cybernetics, vol. 44, no. 11, pp. 1522-1533, 2014
- [7] Monwar M., Gavrilova M., Multimodal Biometric System using Rank-Level Fusion Approach, IEEE Trans. on Systems, Man, and Cybernetics, vol. 39, no. 4, pp. 867-878, 2009
- [8] Rahman M.W., Gavrilova M., Kinect Gait Skeletal Joint Feature Based Person Identification, 16th Int. Conference on Cognitive Informatics & Cognitive Computing (ICCI), IEEE, pp. 423-430, 2017
- [9] Shakhnarovich G., Darrel T., On Probabilistic Combination of Face and Gait Cues for Identification, 5th International Conference on Automatic Face and Gesture Recognition, IEEE, pp. 1-6, 2002
- [10] Rahman M.W., Zohra F.T., Gavrilova M., Rank Level Fusion for Kinect Gait and Face Biometric Identification, Symposium Series on Computational Intelligence (SSCI), IEEE, pp. 2218-2224, 2017
- [11] Kale A., Roychowdhury A., Chellappa R., Fusion of Gait and Face for Human Identification, International Conference on Acoustics, Speech and Signal Processing, IEEE, pp. 1-5, 2004
- [12] Ranganath T., Lee T., Sanei S., Fusion of Chaotic Measure into a New Hybrid Face-Gait System for Human Recognition, 18th International Conference on Pattern Recognition, IEEE, vol. 4, pp. 541-544, 2006
- [13] Zhou X., Bahanu B., Feature Fusion of Face and Gait for Human Recognition at a Distance in Video, 18th International Conference on Pattern Recognition (ICPR), IEEE, vol. 4, pp. 529-532, 2006
- [14] Zhou X., Bahanu B., Integrating Face and Gait for Human Recognition at a Distance in Video, IEEE Trans. on Systems, Man, and Cybernetics, vol. 37, no. 5, pp. 119-1137, 2007
- [15] Geng X. et.al. Adaptive Fusion of Gait and Face for Human Identification in Video, Workshop on Application of Computer Vision, IEEE, pp. 1-6, 2008
- [16] Hossain E., Chetty G., Multimodal Face-Gait Fusion for Biometric Person Authentication, International Conference on Embedded and Ubiquitous Computing (EUC), IEEE, pp. 332-337, 2011
- [17] Hoffman M. et.al, Combined face and gait recognition using alpha matte preprocessing, 5th International Conference on Biometric (ICB), IEEE, pp. 390-395, 2012
- [18] Almohammad M., Salma G., Mahmoud T., Human identification system based on feature level fusion using face and gait biometrics, International Conference on Engineering and Technology (ICET), IEEE, pp. 1-5, 2012

- [19] Xing X., Wang K., Lv Z., Fusion of Gait and Facial Features using Coupled Projections for People Identification at a Distance, *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2349-2353, 2015
- [20] Zhang D. et.al., Ethnicity classification based on fusion of face and gait, *International Conference on Biometric (ICB)*, IEEE, pp. 384-389, 2012
- [21] Zhang D., Wang Y., Gender recognition based on fusion on face and gait information, *International Conference on Machine Learning and Cybernetics*, IEEE, vol. 1, pp. 62-67, 2008
- [22] Ahmed F., Paul P.P., Gavrilova M., DTW based Kernel and Rank Level Fusion for 3D Gait Recognition using Kinect, *The Visual Computer*, vol. 31, no 6, pp. 915-924, 2015
- [23] Vondrick C. et. al. HOGgles: Visualizing Object Detection Features, *International Journal on Computer Vision*, pp. 1-9, 2013
- [24] <http://www.upcv.upatras.gr/personal/kastaniotis/datasets.html> Last accessed: September 20, 2017
- [25] Min R., Kose N., Dugelay, J.L., KinectFaceDB: A Kinect database for face recognition, *IEEE Transactions on Systems, Man and Cybernetics: Systems*, vol. 44, no.11, pp. 1534-1548, 2014
- [26] Paul P.P., Gavrilova M., Alhadj R., Decision Fusion for Multimodal Biometrics Using Social Network Analysis, *Transactions on Systems, Man, and Cybernetics Systems*, IEEE, vol. 44, no. 11, pp. 1522-1533, 2014
- [27] Xing X., Wang K., Lv Z., Fusion of Gait and Facial Features using Coupled Projections for People Identification at a Distance, *Signal Processing Letters*, IEEE, vol. 22, no. 12, pp. 2349-2353, 2015



**Marina L. Gavrilova** is a Full Professor in the Department of Computer Science, University of Calgary, a Head of the Biometric Technologies Laboratory and a Board Member of ISPIA. She has Diploma with Honors from Lomonosov Moscow State University (1993) and Ph D from the University of Calgary (1999). Her publications include

over 200 journal and conference papers, 3 books and 20 book chapters in the areas of image processing, pattern recognition, machine learning, biometric and online security. She is currently serving as a Founding Editor-in-Chief of LNCS Transactions on Computational Science Journal, Springer. Dr. Gavrilova is on the Editorial Boards of the Visual Computer, International Journal of Biometrics, International Journal of Cognitive Biometrics, International Journal of Cybersecurity, and five other journals. She has given over 50 keynotes, invited lectures and tutorials at major scientific gatherings and industry research centers, including at Stanford University, SERIAS C.



**Wasiu Rahman** received his M Sc degree from the Computer Science Department, University of Calgary, in 2018. His thesis was on gait and EEG multi-modal fusion, conducted at the Biometric Technologies laboratory. He is currently employed by the University of Calgary.



**Fatema Tut Zohra** received M Sc degree from the Computer Science Department, University of Calgary, in 2018. She conducted research on biometric data quality at the Biometric Technologies laboratory. She is currently employed as a data scientist by Farmers Edge, Calgary.